

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/12, 15/10, 15/62, C07K 14/72, 16/00, C12Q 1/68, G01N 33/577</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 99/11785</b> <b>(43) International Publication Date:</b> 11 March 1999 (11.03.99)
<b>(21) International Application Number:</b> PCT/GB98/02630 <b>(22) International Filing Date:</b> 2 September 1998 (02.09.98) <b>(30) Priority Data:</b> 9718455.0 2 September 1997 (02.09.97) GB <b>(71) Applicant (for all designated States except US):</b> ROWETT RESEARCH SERVICES LIMITED [GB/GB]; Greenburn Road, Bucksburn, Aberdeen AB21 9SB (GB). <b>(72) Inventor; and</b> <b>(75) Inventor/Applicant (for US only):</b> MCGREGOR, Duncan [GB/GB]; 6 Balcairn Cottages, Oldmeldrum, Aberdeenshire AB51 0EU (GB). <b>(74) Agent:</b> MURGITROYD & COMPANY; 373 Scotland Street, Glasgow G5 8QA (GB).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> CHIMERIC BINDING PEPTIDE LIBRARY SCREENING METHOD  <b>(57) Abstract</b>  There is described a method of isolating nucleotide sequences encoding target peptides from DNA libraries using DNA binding proteins to link the peptide to the sequence which encodes it. DNA libraries are prepared from cells encoding the protein of interest, or from synthetic DNA, and inserted into, or adjacent to, a DNA binding protein in an expression vector to create a chimeric fusion protein. Incorporation of the vector DNA into a carrier package, during expression of the chimeric fusion protein, results in the production of a peptide display carrier package (PDCP) displaying the DNA-bound fusion protein on the external surface of the carrier package. Employment of affinity purification techniques results in the PDCP particles containing sequences encoding the desired peptide to be selected and the desired nucleotide sequences obtained therefrom.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

1      Chimeric binding peptide library screening method

2  
3      The present invention relates generally to methods for  
4      screening nucleotide libraries for sequences that  
5      encode peptides of interest.

6  
7      Isolating an unknown gene which encodes a desired  
8      peptide from a recombinant DNA library can be a  
9      difficult task. The use of hybridisation probes may  
10     facilitate the process, but their use is generally  
11     dependent on knowing at least a portion of the sequence  
12     of the gene which encodes the protein. When the  
13     sequence is not known, DNA libraries can be expressed  
14     in an expression vector, and antibodies have been used  
15     to screen for plaques or colonies displaying the  
16     desired protein antigen. This procedure has been useful  
17     in screening small libraries, but rarely occurring  
18     sequences which are represented in less than about 1 in  
19      $10^5$  clones (as is the case with rarely occurring cDNA  
20     molecules or synthetic peptides) can be easily missed,  
21     making screening libraries larger than  $10^6$  clones at  
22     best laborious and difficult. Methods designed to  
23     address the isolation of rarely occurring sequences by  
24     screening libraries of  $10^6$  clones have been developed  
25     and include phage display methods and LacI fusion phage

1 display, discussed in more detail below.

2  
3 Phage display methods. Members of DNA libraries which  
4 are fused to the N-terminal end of filamentous  
5 bacteriophage pIII and pVIII coat proteins have been  
6 expressed from an expression vector resulting in the  
7 display of foreign peptides on the surface of the phage  
8 particle with the DNA encoding the fusion protein  
9 packaged in the phage particle (Smith G. P., 1985,  
10 Science 228: 1315-1317). The expression vector can be  
11 the bacteriophage genome itself, or a phagemid vector,  
12 into which a bacteriophage coat protein has been  
13 cloned. In the latter case, the host bacterium,  
14 containing the phagemid vector, must be co-infected  
15 with autonomously replicating bacteriophage, termed  
16 helper phage, to provide the full complement of  
17 proteins necessary to produce mature phage particles.  
18 The helper phage normally has a genetic defect in the  
19 origin of replication which results in the preferential  
20 packaging of the phagemid genome. Expression of the  
21 fusion protein following helper phage infection, allows  
22 incorporation of both fusion protein and wild type coat  
23 protein into the phage particle during assembly.  
24 Libraries of fusion proteins incorporated into phage,  
25 can then be selected for binding members against  
26 targets of interest (ligands). Bound phage can then be  
27 allowed to reinfect *Escherichia coli* (*E. coli*) bacteria  
28 and then amplified and the selection repeated,  
29 resulting in the enrichment of binding members  
30 (Parmley, S. F., & Smith, G. P. 1988., Gene 73: 305-  
31 318; Barrett R. W. et al., 1992, Analytical  
32 Biochemistry 204: 357-364 Williamson et al., Proc.  
33 Natl. Acad. Sci. USA, 90: 4141-4145; Marks et al.,  
34 1991, J. Mol. Biol. 222: 581-597).

35  
36 Several publications describe this method. For example,



1 US Patent No 5,403,484 describes production of a  
2 chimeric protein formed from the viral coat protein and  
3 the peptide of interest. In this method at least a  
4 functional portion of a viral coat protein is required  
5 to cause display of the chimeric protein or a processed  
6 form thereof on the outer surface of the virus. In  
7 addition, US Patent No 5,571,698 describes a method for  
8 obtaining a nucleic acid encoding a binding protein, a  
9 key component of which comprises preparing a population  
10 of amplifiable genetic packages which have a  
11 genetically determined outer surface protein, to cause  
12 the display of the potential binding domain on the  
13 outer surface of the genetic package. The genetic  
14 packages are selected from the group consisting of  
15 cells, spores and viruses. For example when the  
16 genetic package is a bacterial cell, the outer surface  
17 transport signal is derived from a bacterial outer  
18 surface protein, and when the genetic package is a  
19 filamentous bacteriophage, the outer surface transport  
20 signal is provided by the gene pIII (minor coat  
21 protein) or pVIII (major coat protein) of the  
22 filamentous phage.

23  
24 WO-A-92/01047 and WO-A-92/20791 describe methods for  
25 producing multimeric specific binding pairs, by  
26 expressing a first polypeptide chain fused to a viral  
27 coat protein, such as the gene pIII protein, of a  
28 secreted replicable genetic display package (RGDP)  
29 which displays a polypeptide at the surface of the  
30 package, and expressing a second polypeptide chain of  
31 the multimer, and allowing the two chains to come  
32 together as part of the RGDP.

33  
34 LacI fusion plasmid display. This method is based on  
35 the DNA binding ability of the lac repressor. Libraries  
36 of random peptides are fused to the lacI repressor

1 protein, normally to the C-terminal end, through  
2 expression from a plasmid vector carrying the fusion  
3 gene. Linkage of the LacI-peptide fusion to its  
4 encoding DNA occurs via the lacO sequences on the  
5 plasmid, forming a stable peptide-LacI-peptide complex.  
6 These complexes are released from their host bacteria  
7 by cell lysis, and peptides of interest isolated by  
8 affinity purification on an immobilised target. The  
9 plasmids thus isolated can then be reintroduced into *E.*  
10 *coli* by electroporation to amplify the selected  
11 population for additional rounds of screening (Cull, M.  
12 G. et al. 1992. Proc. Natl. Acad. Sci. U.S.A. 89:1865-  
13 1869).

14  
15 US Patent No 5498530 describes a method for  
16 constructing a library of random peptides fused to a  
17 DNA binding protein in appropriate host cells and  
18 culturing the host cells under conditions suitable for  
19 expression of the fusion proteins intra-cellularly, in  
20 the cytoplasm of the host cells. This method also  
21 teaches that the random peptide is located at the  
22 carboxy terminus of the fusion protein and that the  
23 fusion protein-DNA complex is released from the host  
24 cell by cell lysis. No method is described for the  
25 protection of the DNA from degradation once released  
26 from the lysed cell. Several DNA binding proteins are  
27 claimed but no examples are shown except lacI.

28  
29 There remains a need for methods of constructing  
30 peptide libraries in addition to the methods described  
31 above. For instance, the above methods do not permit  
32 production of secreted peptides with a free carboxy  
33 terminus. The present invention describes an  
34 alternative method for isolating peptides of interest  
35 from libraries and has significant advantages over the  
36 prior art methods.

1 In general terms, the present invention provides a  
2 method for screening a nucleotide library (usually a  
3 DNA library) for a nucleotide sequence which encodes a  
4 target peptide of interest. The method involves  
5 physically linking each peptide to a polynucleotide  
6 including the specific nucleotide sequence encoding  
7 that peptide. Linkage of a peptide to its encoding  
8 nucleotide sequence is achieved via linkage of the  
9 peptide to a nucleotide binding domain. A bifunctional  
10 chimeric protein with a nucleotide binding domain and a  
11 library member or target peptide (preferably with a  
12 function of interest) is thus obtained. The peptide of  
13 interest is bound to the polynucleotide encoding that  
14 peptide via the nucleotide binding domain of the  
15 chimeric protein.

16  
17 The polynucleotide-chimeric protein complex is then  
18 incorporated within a peptide display carrier package  
19 (PDCP), protecting the polynucleotide from subsequent  
20 degradation, while displaying the target peptide  
21 portion on the outer surface of the peptide display  
22 carrier package (PDCP).

23  
24 Thus, in one aspect, the present invention provides a  
25 peptide display carrier package (PDCP), said package  
26 comprising a polynucleotide-chimeric protein complex  
27 wherein the chimeric protein has a nucleotide binding  
28 portion and a target peptide portion, wherein said  
29 polynucleotide comprises a nucleotide sequence motif  
30 which is specifically bound by said nucleotide binding  
31 portion, and wherein at least the chimeric protein  
32 encoding portion of the polynucleotide not bound by the  
33 nucleotide binding portion of the chimeric protein is  
34 protected.

35  
36 In one embodiment the polynucleotide is protected by a

1 protein which binds non-specifically to naked  
2 polynucleotide. Examples include viral coat proteins,  
3 many of which are well-known in the art. Where the  
4 chosen viral coat protein requires an initiation  
5 sequence to commence general binding to the  
6 polynucleotide, this will be provided on the  
7 polynucleotide at appropriate location(s). A preferred  
8 coat protein is coat protein from a bacteriophage,  
9 especially M13.

10  
11 Generally, the nucleic binding portion of the chimeric  
12 protein is selected for its specificity for the  
13 nucleotide sequence motif present in the recombinant  
14 polynucleotide encoding the chimeric protein.

15  
16 Optionally, the nucleotide sequence motif may be an  
17 integral part of the protein encoding region of the  
18 polynucleotide. Alternatively, and more usually, the  
19 motif may be present in a non-coding region of the  
20 polynucleotide. For the purposes of this invention,  
21 all that is required is for the motif to be located on  
22 the polynucleotide such that the nucleotide binding  
23 portion of the chimeric protein is able to recognise  
24 and bind to it. Desirably the polynucleotide-chimeric  
25 protein complex has a dissociation constant of at least  
26 one hour.

27  
28 Optionally, the recombinant polynucleotide may comprise  
29 two or more nucleotide sequence motifs, each of which  
30 will be bound by a chimeric protein molecule.  
31 Preferably, the motifs are positioned along the length  
32 of the polynucleotide to avoid steric hindrance between  
33 the bound chimeric proteins.

34  
35 Preferably, the nucleotide sequence motif is not  
36 affected by the presence of additional nucleotide

1 sequence (e.g. encoding sequence) at its 5' and/or 3'  
2 ends. Thus the chimeric fusion protein may include a  
3 target peptide portion at its N terminal end, at its C  
4 terminal end or may include two target peptide portions  
5 (which may be the same or different) at each end of the  
6 nucleotide binding portion, ie at both the N and C  
7 terminal ends of the chimeric protein. For example one  
8 target peptide may be an antibody of known specificity  
9 and the other peptide may be a peptide of potential  
10 interest.

11  
12 Desirably the target peptide portion of the chimeric  
13 protein is displayed externally on the peptide display  
14 carrier package, and is thus available for detection,  
15 reaction and/or binding.

16  
17 In more detail the PDCP may be composed two distinct  
18 elements:

19 a. A polynucleotide-chimeric protein complex. This  
20 links the displayed target peptide portion to the  
21 polynucleotide encoding that peptide portion  
22 through a specific polynucleotide binding portion.  
23 The nucleotide sequence encoding the chimeric  
24 protein, and the specific nucleotide sequence  
25 motif recognised by the nucleotide binding portion  
26 of the chimeric protein must be present on a  
27 segment of polynucleotide which can be  
28 incorporated into the PDCP; and

29 b. A protective coat. This may be supplied by a  
30 replicable carrier or helper package capable of  
31 independent existence. Alternatively, a coat  
32 protein could be encoded by the recombinant  
33 polynucleotide of the invention. The protective  
34 coat for the polynucleotide-chimeric protein  
35 complex may be composed of a biological material  
36 such as protein or lipid, but the protective coat

1 is not required for linking the target peptide to  
2 the polynucleotide encoding that peptide. The  
3 protective coat must allow the display of the  
4 target peptide portion of the chimeric protein on  
5 its outer surface. The carrier or helper package  
6 may also provide the mechanism for releasing the  
7 intact PDCP from host cells when so required. By  
8 way of example, when a bacteriophage is the  
9 replicable carrier package, a protein coat of the  
10 bacteriophage surrounds the polynucleotide-  
11 chimeric protein complex to form the PDCP, which  
12 is then extruded from the host bacterial cell.

13  
14 The invention described herein demonstrates that  
15 peptides fused to a nucleotide binding domain can be  
16 displayed externally, even through a bacteriophage  
17 carrier package protein coat, while still bound to the  
18 polynucleotide encoding the displayed peptide.

19  
20 The present invention also provides a recombinant  
21 polynucleotide comprising a nucleotide sequence  
22 encoding a chimeric protein having a nucleotide binding  
23 portion operably linked to a target peptide portion,  
24 wherein said polynucleotide includes a specific  
25 nucleotide sequence motif which is bound by the  
26 nucleotide binding portion of said chimeric protein and  
27 further encoding a non-sequence-specific nucleotide  
28 binding protein.

29  
30 Desirably, the recombinant polynucleotide is a  
31 recombinant expression system, able to express the  
32 chimeric protein when placed in a suitable environment,  
33 for example a compatible host cell. After its  
34 expression, the chimeric protein binds to the specific  
35 nucleotide sequence (motif) present in the  
36 polynucleotide comprising the nucleotide sequence

1 encoding the chimeric protein.

2

3 Optionally there may be a linker sequence located  
4 between the nucleotide sequence encoding the nucleotide  
5 binding portion and the polynucleotide inserted into  
6 the restriction enzyme site of the construct.

7

8 Desirably the nucleotide binding portion is a DNA  
9 binding domain of an oestrogen or progesterone  
10 receptor, or a functional equivalent thereof. Examples  
11 of sequences encoding such nucleotide binding portions  
12 are set out in SEQ ID Nos 11 and 13.

13

14 The term "expression system" is used herein to refer to  
15 a genetic sequence which includes a protein-encoding  
16 region and is operably linked to all of the genetic  
17 signals necessary to achieve expression of that region.  
18 Optionally, the expression system may also include  
19 regulatory elements, such as a promoter or enhancer to  
20 increase transcription and/or translation of the  
21 protein encoding region or to provide control over  
22 expression. The regulatory elements may be located  
23 upstream or downstream of the protein encoding region  
24 or within the protein encoding region itself. Where  
25 two or more distinct protein encoding regions are  
26 present these may use common regulatory element(s) or  
27 have separate regulatory element(s).

28

29 Generally, the recombinant polynucleotide described  
30 above will be DNA. Where the expression system is  
31 based upon an M13 vector, usually the polynucleotide  
32 binding portion of the expressed chimeric portion will  
33 be single-stranded DNA. However, other vector systems  
34 may be used and the nucleotide binding portion may be  
35 selected to bind preferentially to double-stranded DNA  
36 or to double or single-stranded RNA, as convenient.

1     Additionally the present invention provides a vector  
2     containing such a recombinant expression system and  
3     host cells transformed with such a recombinant  
4     expression system (optionally in the form of a vector).  
5

6     Whilst the recombinant polynucleotide described above  
7     forms an important part of the present invention, we  
8     are also concerned with the ability to screen large  
9     (e.g. of at least  $10^5$  members, for example  $10^6$  or even  
10     $10^7$  members) libraries of genetic material. One of the  
11    prime considerations therefore is the provision of a  
12    recombinant genetic construct into which each member of  
13    said library can individually be incorporated to form  
14    the recombinant polynucleotide described above and to  
15    express the chimeric protein thereby encoded (the  
16    target peptide of which is encoded by the nucleotide  
17    library member incorporated into the construct).  
18

19    Thus viewed in a further aspect the present invention  
20    provides a genetic construct or set of genetic  
21    constructs comprising a polynucleotide having a  
22    sequence which includes:  
23

- 24    i)    a sequence encoding a nucleotide binding portion  
25        able to recognise and bind to a specific sequence  
26        motif;
- 27    ii)   the sequence motif recognised and bound by the  
28        nucleotide binding portion encoded by (i);
- 29    iii) a restriction enzyme site which permits insertion  
30        of a polynucleotide, said site being designed to  
31        operably link said polynucleotide to the sequence  
32        encoding the nucleotide binding portion so that  
33        expression of the operably linked polynucleotide  
34        sequences yields a chimeric protein; and
- 35    iv)   a sequence encoding a nucleotide binding protein  
36        which binds non-specifically to naked



1 polynucleotide.

2

3 Optionally there may be a linker sequence located  
4 between the nucleotide sequence encoding the nucleotide  
5 binding portion and the sequence of the polynucleotide  
6 from the library inserted into the restriction enzyme  
7 site of the construct.

8

9 Desirably the nucleotide binding portion is a DNA  
10 binding domain of an oestrogen or progesterone  
11 receptor, or a functional equivalent thereof. Examples  
12 of sequences encoding such nucleotide binding portions  
13 are set out in SEQ ID Nos 11 and 13.

14

15 Suitable genetic constructs according to the invention  
16 include pDM12, pDM14 and pDM16, deposited at NCIMB on  
17 28 August 1998 under Nos NCIMB 40970, NCIMB 40971 and  
18 NCIMB 40972 respectively.

19

20 It is envisaged that a conventionally produced genetic  
21 library may be exposed to the genetic construct(s)  
22 described above. Thus, each individual member of the  
23 genetic library will be separately incorporated into  
24 the genetic construct and the library will be present  
25 in the form of a library of recombinant polynucleotides  
26 (as described above), usually in the form of vectors,  
27 each recombinant polynucleotide including as library  
28 member.

29

30 Thus, in a further aspect, the present invention  
31 provides a library of recombinant polynucleotides (as  
32 defined above) wherein each polynucleotide includes a  
33 polynucleotide obtained from a genetic library and  
34 which encodes the target peptide portion of the  
35 chimeric protein expressed by the recombinant  
36 polynucleotide.

1     Optionally, the chimeric protein may further include a  
2     linker sequence located between the nucleotide binding  
3     portion and the target peptide portion. The linker  
4     sequence will reduce steric interference between the  
5     two portions of the protein. Desirably the linker  
6     sequence exhibits a degree of flexibility.

7  
8     Also disclosed are methods for constructing and  
9     screening libraries of PDCP particles, displaying many  
10    different peptides, allowing the isolation and  
11    identification of particular peptides by means of  
12    affinity techniques relying on the binding activity of  
13    the peptide of interest. The resulting polynucleotide  
14    sequences can therefore be more readily identified, re-  
15    cloned and expressed.

16  
17    A method of constructing a genetic library, said method  
18    comprising:

- 19  
20    a)    constructing multiple copies of a recombinant  
21          vector comprising a polynucleotide sequence which  
22          encodes a nucleotide binding portion able to  
23          recognise and bind to a specific sequence motif  
24          (and optionally also including the specific  
25          sequence motif);  
26  
27    b)    operably linking each said vector to a  
28          polynucleotide encoding a target polypeptide, such  
29          that expression of said operably linked vector  
30          results in expression of a chimeric protein  
31          comprising said target peptide and said nucleotide  
32          binding portions; wherein said multiple copies of  
33          said operably linked vectors collectively express  
34          a library of target peptide portions;  
35  
36    c)    transforming host cells with the vectors of step

- 1           b) ;
- 2
- 3       d)    culturing the host cells of step c) under
- 4            conditions suitable for expression of said
- 5            chimeric protein;
- 6
- 7       e)    providing a recombinant polynucleotide comprising
- 8            the nucleotide sequence motif specifically
- 9            recognised by the nucleotide binding portion and
- 10           exposing this polynucleotide to the chimeric
- 11           protein of step d) to yield a polynucleotide-
- 12           chimeric protein complex; and
- 13
- 14       f)    causing production of a non-sequence-specific
- 15            moiety able to bind to the non-protected portion
- 16            of the polynucleotide encoding the chimeric
- 17            protein to form a peptide display carrier package.

18

19       The present invention further provides a method of

20       screening a genetic library, said method comprising:

- 21
- 22       a)    exposing the polynucleotide members of said
- 23            library to multiple copies of a genetic construct
- 24            comprising a nucleotide sequence encoding a
- 25            nucleotide binding portion able to recognise and
- 26            bind to a specific sequence motif, under
- 27            conditions suitable for the polynucleotides of
- 28            said library each to be individually ligated into
- 29            one copy of said genetic construct, to create a
- 30            library of recombinant polynucleotides;
- 31
- 32       b)    exposing said recombinant polynucleotides to a
- 33            population of host cells, under conditions
- 34            suitable for transformation of said host cells by
- 35            said recombinant polynucleotides;

36

- 1 c) selecting for transformed host cells;
- 2
- 3 d) exposing said transformed host cells to conditions
- 4 suitable for expression of said recombinant
- 5 polynucleotide to yield a chimeric protein; and
- 6
- 7 e) providing a recombinant polynucleotide comprising
- 8 the nucleotide sequence motif specifically
- 9 recognised by the nucleotide binding portion and
- 10 exposing this polynucleotide to the chimeric
- 11 protein of step d) to yield a polynucleotide-
- 12 chimeric protein complex;
- 13
- 14 f) protecting any exposed portions of the
- 15 polynucleotide in the complex of step e) to form a
- 16 peptide display carrier package; and
- 17
- 18 g) screening said peptide display carrier package to
- 19 select only those packages displaying a target
- 20 peptide portion having the characteristics
- 21 required.

22  
23 Desirably in step a) the genetic construct is pDM12,  
24 pDM14 or pDM16.

25  
26 Desirably in step f) the peptide display package  
27 carrier is extruded from the transformed host cell  
28 without lysis of the host cell.

29  
30 Generally the transformed host cells will be plated out  
31 or otherwise divided into single colonies following  
32 transformation and prior to expression of the chimeric  
33 protein.

34  
35 The screening step g) described above may look for a  
36 particular target peptide either on the basis of

1 function (e.g. enzymic activity) or structure (e.g.  
2 binding to a specific antibody). Once the peptide  
3 display carrier package is observed to include a target  
4 peptide with the desired characteristics, the  
5 polynucleotide portion thereof (which of course encodes  
6 the chimeric protein itself) can be amplified, cloned  
7 and otherwise manipulated using standard genetic  
8 engineering techniques.

9  
10 The current invention differs from the prior art  
11 teaching of the previous disclosures US Patent No  
12 5,403,484 and US Patent No 5,571,698, as the invention  
13 does not require outer surface transport signals, or  
14 functional portions of viral coat proteins, to enable  
15 the display of chimeric binding proteins on the outer  
16 surface of the viral particle or genetic package.

17  
18 The current invention also differs from the teaching of  
19 WO-A-92/01047 and WO-A-92/20791, as no component of a  
20 secreted replicable genetic display package, or viral  
21 coat protein is required, to enable display of the  
22 target peptide on the outer surface of the viral  
23 particle.

24  
25 The current invention differs from the teaching of US  
26 Patent No 5498530, as it enables the display of  
27 chimeric proteins, linked to the polynucleotide  
28 encoding the chimeric protein, extra-cellularly, not in  
29 the cytoplasm of a host cell. In the current invention  
30 the chimeric proteins are presented on the outer  
31 surface of a peptide display carrier package (PDCP)  
32 which protects the DNA encoding the chimeric protein,  
33 and does not require cell lysis to obtain access to the  
34 chimeric protein-DNA complex. Finally, the current  
35 invention does not rely upon the lacI DNA binding  
36 protein to form the chimeric protein-DNA complex.

1 In one embodiment of the invention, the nucleotide  
2 binding portion of the chimeric protein comprises a DNA  
3 binding domain from one or more of the nuclear steroid  
4 receptor family of proteins, or a functional equivalent  
5 of such a domain. Particular examples include (but are  
6 not limited to) a DNA binding domain of the oestrogen  
7 receptor or the progesterone receptor, or functional  
8 equivalents thereof. These domains can recognise  
9 specific DNA sequences, termed hormone response  
10 elements (HRE), which can be bound as both double and  
11 single-stranded DNA. The DNA binding domain of such  
12 nuclear steroid receptor proteins is preferred.

13  
14 The oestrogen receptor is especially referred to below  
15 by way of example, for convenience since:

16 (a) The oestrogen receptor is a large multifunctional  
17 polypeptide of 595 amino acids which functions in the  
18 cytoplasm and nucleus of eukaryotic cells (Green et  
19 al., 1986, Science 231: 1150-1154). A minimal high  
20 affinity DNA binding domain (DBD) has been defined  
21 between amino acids 176 and 282 (Mader et al., 1993,  
22 Nucleic Acids Res. 21: 1125-1132). The functioning of  
23 this domain (i.e. DNA binding) is not inhibited by the  
24 presence of non-DNA binding domains at both the N and C  
25 terminal ends of this domain, in the full length  
26 protein.

27  
28 (b) The oestrogen receptor DNA binding domain fragment  
29 (amino acids 176-282) has been expressed in *E. coli* and  
30 shown to bind to the specific double stranded DNA  
31 oestrogen receptor target HRE nucleotide sequence, as a  
32 dimer with a similar affinity (0.5nM) to the parent  
33 molecule (Murdoch et al. 1990, Biochemistry 29: 8377-  
34 8385; Mader et al., 1993, Nucleic Acids Research 21:  
35 1125-1132). DBD dimerization on the surface of the PDCP  
36 should result in two peptides displayed per particle.

1 This bivalent display can aid in the isolation of low  
2 affinity peptides and peptides that are required to  
3 form a bivalent conformation in order to bind to a  
4 particular target, or activate a target receptor. The  
5 oestrogen receptor is capable of binding to its 38 base  
6 pair target HRE sequence, consensus sequence:

7  
8 1) 5'-TCAGGTCAGAGTGACCTGAGCTAAAATAACACATTCAG-3'  
9 ("minus strand") SEQ ID No 1, and

10  
11 2) 3'-AGTCCAGTCTCACTGGACTCGATTTTATTGTGTAAGTC-5'  
12 ("plus strand") SEQ ID No 2,

13  
14 with high affinity and specificity, under the salt and  
15 pH conditions normally required for selection of  
16 binding peptides. Moreover, binding affinity is  
17 increased 60-fold for the single-stranded coding, or  
18 "plus", strand (i.e. SEQ ID No 2) of the HRE nucleotide  
19 sequence over the double stranded form of the specific  
20 target nucleotide sequence (Peale et al. 1988, Proc.  
21 Natl. Acad. Sci. USA 85: 1038-1042; Lannigan & Notides,  
22 1989, Proc. Natl. Acad. Sci. USA 86: 863-867).

23  
24 In an embodiment of the invention where the DNA binding  
25 component of the peptide display carrier package is the  
26 oestrogen receptor, the nucleotide (DNA) binding  
27 portion contains a minimum sequence of amino acids 176-  
28 282 of the oestrogen receptor protein. In addition, the  
29 consensus oestrogen receptor target HRE sequence is  
30 cloned in such a way that if single stranded DNA can be  
31 produced then the coding, or "plus", strand of the  
32 oestrogen receptor HRE nucleotide sequence is  
33 incorporated into single-stranded DNA. An example of a  
34 vector suitable for this purpose is pUC119 (see Viera  
35 et al., Methods in Enzymology, Vol 153, pages 3-11,  
36 1987).

1 In a preferred embodiment of the invention a peptide  
2 display carrier package (PDCP) can be assembled when a  
3 bacterial host cell is transformed with a bacteriophage  
4 vector, which vector comprises a recombinant  
5 polynucleotide as described above. The expression  
6 vector will also comprise the specific nucleotide motif  
7 that can be bound by the nucleotide binding portion of  
8 the chimeric protein. Expression of recombinant  
9 polynucleotide results in the production of the  
10 chimeric protein which comprises the target peptide and  
11 the nucleotide binding portion. The host cell is grown  
12 under conditions suitable for chimeric protein  
13 expression and assembly of the bacteriophage particles,  
14 and the association of the chimeric protein with the  
15 specific nucleotide sequence in the expression vector.

16  
17 In this embodiment, since the vector is a  
18 bacteriophage, which replicates to produce a single-  
19 stranded DNA, the nucleotide binding portion preferably  
20 has an affinity for single-stranded DNA. Incorporation  
21 of the vector single-stranded DNA-chimeric protein  
22 complex into bacteriophage particles results in the  
23 assembly of the peptide display carrier package (PDCP),  
24 and display of the target peptide on the outer surface  
25 of the PDCP.

26  
27 In this embodiment both of the required elements for  
28 producing peptide display carrier packages are  
29 contained on the same vector. Incorporation of the DNA-  
30 chimeric protein complex into a peptide display carrier  
31 package (PDCP) is preferred as DNA degradation is  
32 prevented, large numbers of PDCPs are produced per host  
33 cell, and the PDCPs are easily separated from the host  
34 cell without recourse to cell lysis.

35  
36 In a more preferred embodiment, the vector of the is a



1 phagemid vector (for example pUC119) where expression  
2 of the chimeric protein is controlled by an inducible  
3 promoter. In this embodiment the PDCP can only be  
4 assembled following infection of the host cell with  
5 both phagemid vector and helper phage. The transfected  
6 host cell is then cultivated under conditions suitable  
7 for chimeric protein expression and assembly of the  
8 bacteriophage particles.

9  
10 In this embodiment the elements of the PDCP are  
11 provided by two separate vectors. The phagemid derived  
12 PDCP is superior to phagemid derived display packages  
13 disclosed in WO-A-92/01047 where a proportion of  
14 packages displaying bacteriophage coat protein fusion  
15 proteins will contain the helper phage DNA, not the  
16 fusion protein DNA sequence. In the current invention,  
17 a PDCP can display the chimeric fusion protein only  
18 when the package contains the specific nucleotide motif  
19 recognised by the nucleotide binding portion. In most  
20 embodiments this sequence will be present on the same  
21 DNA segment that encodes the fusion protein. In  
22 addition, the prior art acknowledges that when mutant  
23 and wild type proteins are co-expressed in the same  
24 bacterial cell, the wild type protein is produced  
25 preferentially. Thus, when the wild type helper phage,  
26 phage display system of WO-A-92/01047 is used, both  
27 wild type gene pIII and target peptide-gene pIII  
28 chimeric proteins are produced in the same cell. The  
29 result of this is that the wild type gene pIII protein  
30 is preferentially packaged into bacteriophage  
31 particles, over the chimeric protein. In the current  
32 invention, there is no competition with wild type  
33 bacteriophage coat proteins for packaging.

34  
35 Desirably the target peptide is displayed in a location  
36 exposed to the external environment of the PDCP, after

1 the PDCP particle has been released from the host cell  
2 without recourse to cell lysis. The target peptide is  
3 then accessible for binding to its ligand. Thus, the  
4 target peptide may be located at or near the N-terminus  
5 or the C-terminus of a nucleotide binding domain, for  
6 example the DNA binding domain of the oestrogen  
7 receptor.

8

9 The present invention also provides a method for  
10 screening a DNA library expressing one or more  
11 polypeptide chains that are processed, folded and  
12 assembled in the periplasmic space to achieve  
13 biological activity. The PDCP may be assembled by the  
14 following steps:

15

16 (a) Construction of N- or C-terminal DBD chimeric  
17 protein fusions in a phagemid vector.

18 (i) When the target peptide is located at the N-  
19 terminus of the nucleotide binding portion, a library  
20 of DNA sequences each encoding a potential target  
21 peptide is cloned into an appropriate location of an  
22 expression vector (i.e. behind an appropriate promoter  
23 and translation sequences and a sequence encoding a  
24 signal peptide leader directing transport of the  
25 downstream fusion protein to the periplasmic space) and  
26 upstream of the sequence encoding the nucleotide  
27 binding portion. In a preferred embodiment the DNA  
28 sequence(s) of interest may be joined, by a region of  
29 DNA encoding a flexible amino acid linker, to the 5'-  
30 end of an oestrogen receptor DBD.

31 (ii) Alternatively, when the target peptide is  
32 located at the C-terminus of the nucleotide binding  
33 domain, a library of DNA sequences each encoding a  
34 potential target peptide is cloned into the expression  
35 vector so that the nucleotide sequence coding for the  
36 nucleotide binding portion is upstream of the cloned

1 DNA target peptide encoding sequences, said nucleotide  
2 binding portion being positioned behind an appropriate  
3 promoter and translation sequences and a sequence  
4 encoding a signal peptide leader directing transport of  
5 the downstream fusion protein to the periplasmic space.  
6 In a preferred embodiment, DNA sequence(s) of interest  
7 may be joined, by a region of DNA encoding a flexible  
8 amino acid linker oestrogen receptor DBD DNA sequence.

9  
10 Located on the expression vector is the specific HRE  
11 nucleotide sequence recognised, and bound, by the  
12 oestrogen receptor DBD. In order to vary the number of  
13 chimeric proteins displayed on each PDCP particle, this  
14 sequence can be present as one or more copies in the  
15 vector.

16  
17 (b) Incorporation into the PDCP. Non-lytic helper  
18 bacteriophage infects host cells containing the  
19 expression vector. Preferred types of bacteriophage  
20 include the filamentous phage fd, fl and M13. In a  
21 more preferred embodiment the bacteriophage may be  
22 M13K07.

23  
24 The protein(s) of interest are expressed and  
25 transported to the periplasmic space, and the properly  
26 assembled proteins are incorporated into the PDCP  
27 particle by virtue of the high affinity interaction of  
28 the DBD with the specific target nucleotide sequence  
29 present on the phagemid vector DNA which is naturally  
30 packaged into phage particles in a single-stranded  
31 form. The high affinity interaction between the DBD  
32 protein and its specific target nucleotide sequence  
33 prevents displacement by bacteriophage coat proteins  
34 resulting in the incorporation of the protein(s) of  
35 interest onto the surface of the PDCP as it is extruded  
36 from the cell.

1 (c) Selection of the peptide of interest. Particles  
2 which display the peptide of interest are then selected  
3 from the culture by affinity enrichment techniques.  
4 This is accomplished by means of a ligand specific for  
5 the protein of interest, such as an antigen if the  
6 protein of interest is an antibody. The ligand may be  
7 presented on a solid surface such as the surface of an  
8 ELISA plate, or in solution. Repeating the affinity  
9 selection procedure provides an enrichment of clones  
10 encoding the desired sequences, which may then be  
11 isolated for sequencing, further cloning and/or  
12 expression.

13  
14 Numerous types of libraries of peptides fused to the  
15 DBD can be screened under this embodiment including:

16  
17 (i) Random peptide sequences encoded by synthetic  
18 DNA of variable length.

19  
20 (ii) Single-chain Fv antibody fragments. These  
21 consist of the antibody heavy and light chain  
22 variable region domains joined by a flexible  
23 linker peptide to create a single-chain antigen  
24 binding molecule.

25  
26 (iii) Random fragments of naturally occurring  
27 proteins isolated from a cell population  
28 containing an activity of interest.

29  
30 In another embodiment the invention concerns methods  
31 for screening a DNA library whose members require more  
32 than one chain for activity, as required by, for  
33 example, antibody Fab fragments for ligand binding. In  
34 this embodiment heavy or light chain antibody DNA is  
35 joined to a nucleotide sequence encoding a DNA binding  
36 domain of, for example, the oestrogen receptor in a

1 phagemid vector. Typically the antibody DNA library  
2 sequences for either the heavy (VH and CH1) or light  
3 chain (VL and CL) genes are inserted in the 5' region  
4 of the oestrogen receptor DBD DNA, behind an  
5 appropriate promoter and translation sequences and a  
6 sequence encoding a signal peptide leader directing  
7 transport of the downstream fusion protein to the  
8 periplasmic space.

9

10 Thus, a DBD fused to a DNA library member-encoded  
11 protein is produced and assembled in to the viral  
12 particle after infection with bacteriophage. The second  
13 and any subsequent chain(s) are expressed separately  
14 either:

15

16 (a) from the same phagemid vector containing the DBD  
17 and the first polypeptide fusion protein,  
18 or

19

20 (b) from a separate region of DNA which may be present  
21 in the host cell nucleus, or on a plasmid, phagemid or  
22 bacteriophage expression vector that can co-exist, in  
23 the same host cell, with the first expression vector,  
24 so as to be transported to the periplasm where they  
25 assemble with the first chain that is fused to the DBD  
26 protein as it exits the cell. Peptide display carrier  
27 packages (PDCP) which encode the protein of interest  
28 can then be selected by means of a ligand specific for  
29 the protein.

30

31 In yet another embodiment, the invention concerns  
32 screening libraries of bi-functional peptide display  
33 carrier packages where two or more activities of  
34 interest are displayed on each PDCP. In this  
35 embodiment, a first DNA library sequence(s) is inserted  
36 next to a first DNA binding domain (DBD) DNA sequence,

1 for example the oestrogen receptor DBD, in an  
2 appropriate vector, behind an appropriate promoter and  
3 translation sequences and a sequence encoding a signal  
4 peptide leader directing transport of this first  
5 chimeric protein to the periplasmic space. A second  
6 chimeric protein is also produced from the same, or  
7 separate, vector by inserting a second DNA library  
8 sequence(s) next to a second DBD DNA sequence which is  
9 different from the first DBD DNA sequence, for example  
10 the progesterone receptor DBD, behind an appropriate  
11 promoter and translation sequences and a sequence  
12 encoding a signal peptide leader. The first, or only,  
13 vector contains the specific HRE nucleotide sequences  
14 for both oestrogen and progesterone receptors.  
15 Expression of the two chimeric proteins, results in a  
16 PDCP with two different chimeric proteins displayed. As  
17 an example, one chimeric protein could possess a  
18 binding activity for a particular ligand of interest,  
19 while the second chimeric protein could possess an  
20 enzymatic activity. Binding by the PDCP to the ligand  
21 of the first chimeric protein could then be detected by  
22 subsequent incubation with an appropriate substrate for  
23 the second chimeric protein. In an alternative  
24 embodiment a bi-functional PDCP may be created using a  
25 single DBD, by cloning one peptide at the 5'-end of the  
26 DBD, and a second peptide at the 3'-end of the DBD.  
27 Expression of this single bi-functional chimeric  
28 protein results in a PDCP with two different  
29 activities.

30  
31 We have investigated the possibility of screening  
32 libraries of peptides, fused to a DNA binding domain  
33 and displayed on the surface of a display package, for  
34 particular peptides with a biological activity of  
35 interest and recovering the DNA encoding that activity.  
36 Surprisingly, by manipulating the oestrogen receptor

1 DNA binding domain in conjunction with M13  
2 bacteriophage we have been able to construct novel  
3 particles which display large biologically functional  
4 molecules, that allows enrichment of particles with the  
5 desired specificity.

6  
7 The invention described herein provides a significant  
8 breakthrough in DNA library screening technology.

9  
10 The invention will now be further described by  
11 reference to the non-limiting examples and figures  
12 below.

#### 13 14 Description of Figures

15  
16 Figure 1 shows the pDM12 N-terminal fusion oestrogen  
17 receptor DNA binding domain expression vector  
18 nucleotide sequence (SEQ ID No 3), between the HindIII  
19 and EcoRI restriction sites, comprising a pelB leader  
20 secretion sequence (in italics), multiple cloning site  
21 containing SfiI and NotI sites, flexible (glycine)<sub>4</sub>-  
22 serine linker sequence (boxed), a fragment of the  
23 oestrogen receptor gene comprising amino acids 176-282  
24 (SEQ ID No 4) of the full length molecule, and the 38  
25 base pair consensus oestrogen receptor DNA binding  
26 domain HRE sequence.

27  
28 Figure 2 shows the OD<sub>450nm</sub> ELISA data for negative  
29 control M13K07 phage, and single-clone PDCP display  
30 culture supernatants (#1-4, see Example 3) isolated by  
31 selection of the lymphocyte cDNA-pDM12 library against  
32 anti-human immunoglobulin kappa antibody.

33  
34 Figure 3 shows partial DNA (SEQ ID No 5) and amino acid  
35 (SEQ ID No 6) sequence for the human immunoglobulin  
36 kappa constant region (Kabat, E. A. et al., Sequences

1 of Proteins of Immunological Interest. 4<sup>th</sup> edition. U.S.  
2 Department of Health and Human Services. 1987), and  
3 ELISA positive clones #2 (SEQ ID Nos 7 and 8) and #3  
4 (SEQ ID Nos 9 and 10) from Figure 2 which confirms the  
5 presence of human kappa constant region DNA in-frame  
6 with the pelB leader sequence (pelB leader sequence is  
7 underlined, the leader sequence cleavage site is  
8 indicated by an arrow). The differences in the 5'-end  
9 sequence demonstrates that these two clones were  
10 selected independently from the library stock. The PCR  
11 primer sequence is indicated in bold, clone #2 was  
12 originally amplified with CDNAPCRBAK1 and clone #3 was  
13 amplified with CDNAPCRBAK2.

14  
15 Figure 4 shows the pDM14 N-terminal fusion oestrogen  
16 receptor DNA binding domain expression vector  
17 nucleotide sequence (SEQ ID No 11), between the HindIII  
18 and EcoRI restriction sites, comprising a pelB leader  
19 secretion sequence (in italics), multiple cloning site  
20 containing SfiI and NotI sites, flexible (glycine)<sub>4</sub>-  
21 serine linker sequence (boxed), a fragment of the  
22 oestrogen receptor gene comprising amino acids 176-282  
23 (SEQ ID No 12) of the full length molecule, and the two  
24 38 base pair oestrogen receptor DNA binding domain HRE  
25 sequences (HRE 1 and HRE 2).

26  
27 Figure 5 shows the pDM16 C-terminal fusion oestrogen  
28 receptor DNA binding domain expression vector  
29 nucleotide sequence (SEQ ID No 13), between the HindIII  
30 and EcoRI restriction sites, comprising a pelB leader  
31 secretion sequence (in italics), a fragment of the  
32 oestrogen receptor gene comprising amino acids 176-282  
33 (SEQ ID No 14) of the full length molecule, flexible  
34 (glycine)<sub>4</sub>-serine linker sequence (boxed), multiple  
35 cloning site containing SfiI and NotI sites and the 38  
36 base pair oestrogen receptor DNA binding domain HRE



1 sequence.

2

3 Figure 6 shows the OD<sub>450nm</sub> ELISA data for N-cadherin-  
4 pDM16 C-terminal display PDCP binding to anti-pan-  
5 cadherin monoclonal antibody in serial dilution ELISA  
6 as ampicillin resistance units (a.r.u.). Background  
7 binding of negative control M13K07 helper phage is also  
8 shown.

9

10 Figure 7 shows the OD<sub>450nm</sub> ELISA data for *in vivo*  
11 biotinylated PCC-pDM16 C-terminal display PDCP binding  
12 to streptavidin in serial dilution ELISA as ampicillin  
13 resistance units (a.r.u.). Background binding of  
14 negative control M13K07 helper phage is also shown.

15

16 Figure 8 shows the OD<sub>450nm</sub> ELISA data for a human scFv  
17 PDCP isolated from a human scFv PDCP display library  
18 selected against substance P. The PDCP was tested  
19 against streptavidin (1), streptavidin-biotinylated  
20 substance P (2), and streptavidin-biotinylated CGRP  
21 (3), in the presence (B) or absence (A) of free  
22 substance P.

23

24 Figure 9 shows the DNA (SEQ ID Nos 15 and 17) and amino  
25 acid (SEQ ID No 16 and 18) sequence of the substance P  
26 binding scFv isolated from a human scFv PDCP display  
27 library selected against substance P. Heavy chain (SEQ  
28 ID Nos 15 and 16) and light chain (SEQ ID Nos 17 and  
29 18) variable region sequence is shown with the CDRs  
30 underlined and highlighted in bold.

31

## 32 Materials and Methods

33 The following procedures used by the present applicant  
34 are described in Sambrook, J., et al., 1989 *supra*.:  
35 restriction enzyme digestion, ligation, preparation of  
36 electrocompetent cells, electroporation, analysis of

1 restriction enzyme digestion products on agarose gels,  
2 DNA purification using phenol/chloroform, preparation  
3 of 2xTY medium and plates, preparation of ampicillin,  
4 kanamycin, IPTG (Isopropyl  $\beta$ -D-Thiogalactopyranoside)  
5 stock solutions, and preparation of phosphate buffered  
6 saline.

7  
8 Restriction enzymes, T4 DNA ligase and cDNA synthesis  
9 reagents (Superscript plasmid cDNA synthesis kit) were  
10 purchased from Life Technologies Ltd (Paisley,  
11 Scotland, U.K.). Oligonucleotides were obtained from  
12 Cruachem Ltd (Glasgow, Scotland, U.K.), or Genosys  
13 Biotechnologies Ltd (Cambridge, U.K.). Taq DNA  
14 polymerase, Wizard SV plasmid DNA isolation kits,  
15 streptavidin coated magnetic beads and mRNA isolation  
16 reagents (PolyATract 1000) were obtained from Promega  
17 Ltd (Southampton, Hampshire, U.K.). Taqplus DNA  
18 polymerase was obtained from Stratagene Ltd (Cambridge,  
19 U.K.). PBS, BSA, streptavidin, substance P and anti-pan  
20 cadherin antibody were obtained from SIGMA Ltd (Poole,  
21 Dorset, U.K.). Anti-M13-HRP conjugated antibody,  
22 Kanamycin resistant M13K07 helper bacteriophage and  
23 RNAGuard were obtained from Pharmacia Ltd (St. Albans,  
24 Herts, U.K.) and anti-human Ig $\kappa$  antibody from Harlan-  
25 Seralab (Loughborough, Leicestershire, U.K.)  
26 Biotinylated substance P and biotinylated calcitonin  
27 gene related peptide (CGRP) were obtained from  
28 Peninsula Laboratories (St. Helens, Merseyside, U.K.).

29  
30 Specific embodiments of the invention are given below  
31 in Examples 1 to 9.

32

1     **Example 1. Construction of a N-terminal PDCP display**  
2     **phagemid vector pDM12.**

3  
4     The pDM12 vector was constructed by inserting an  
5     oestrogen receptor DNA binding domain, modified by  
6     appropriate PCR primers, into a phagemid vector pDM6.  
7     The pDM6 vector is based on the pUC119 derived phage  
8     display vector pHEN1 (Hoogenboom et al., 1991, Nucleic  
9     Acids Res. 19: 4133-4137). It contains (Gly)<sub>4</sub>Ser linker,  
10    Factor Xa cleavage site, a full length gene III, and  
11    streptavidin tag peptide sequence (Schmidt, T.G. and  
12    Skerra, A., 1993, Protein Engineering 6: 109-122), all  
13    of which can be removed by NotI-EcoRI digestion and  
14    agarose gel electrophoresis, leaving a pelB leader  
15    sequence, SfiI, NcoI and PstI restriction sites  
16    upstream of the digested NotI site. The cloned DNA  
17    binding domain is under the control of the lac promoter  
18    found in pUC119.

19  
20    **Preparation of pDM6**

21  
22    The pDM12 vector was constructed by inserting an  
23    oestrogen receptor DNA binding domain, modified by  
24    appropriate PCR primers, into a phagemid vector pDM6.  
25    The pDM6 vector is based on the gene pIII phage display  
26    vector pHEN1 (Hoogenboom et al., 1991, Nucleic Acids  
27    Res. 19: 4133-4137), itself derived from pUC119 (Viera,  
28    J. and Messing, J., 1987, Methods in Enzymol. 153:  
29    3-11). It was constructed by amplifying the pIII gene  
30    in pHEN1 with two oligonucleotides:

31  
32    PDM6BAK: 5 -TTT TCT GCA GTA ATA GGC GGC CGC AGG GGG AGG  
33    AGG GTC CAT CGA AGG TCG CGA AGC AGA GAC TGT TGA AAG T-3  
34    (SEQ ID No 19) and

35  
36    PDM6FOR: 5 - TTT TGA ATT CTT ATT AAC CAC CGA ACT GCG

1 GGT GAC GCC AAG CGC TTG CGG CCG TTA AGA CTC CTT ATT ACG  
2 CAG-3 (SEQ ID No 20).

3  
4 and cloning the PstI-EcoRI digested PCR product back  
5 into similarly digested pHEN1, thereby removing the  
6 c-myc tag sequence and supE TAG codon from pHEN1. The  
7 pDM6 vector contains a (Gly)<sub>4</sub>Ser linker, Factor Xa  
8 cleavage site, a full length gene III, and streptavidin  
9 tag peptide sequence (Schmidt, T.G. and Skerra, A.,  
10 1993, Protein Engineering 6: 109-122), all of which can  
11 be removed by NotI-EcoRI digestion and agarose gel  
12 electrophoresis, leaving a pelB leader sequence, SfiI,  
13 NcoI and PstI restriction sites upstream of the  
14 digested NotI site. The cloned DNA binding domain is  
15 under the control of the lac promoter found in pUC119.

16  
17 The oestrogen receptor DNA binding domain was isolated  
18 from cDNA prepared from human bone marrow (Clontech,  
19 Palo Alto, California, U.S.A.). cDNA can be prepared by  
20 many procedures well known to those skilled in the art.  
21 As an example, the following method using a Superscript  
22 plasmid cDNA synthesis kit can be used:

23  
24 (a) First strand synthesis.

25  
26 5µg of bone marrow mRNA, in 5µl DEPC-treated water was  
27 thawed on ice and 2µl (50pmol) of cDNA synthesis primer  
28 (5'-AAAAGCGGCCGCACTGGCCTGAGAGA(N)<sub>6</sub>-3') (SEQ ID No 21)  
29 was added to the mRNA and the mixture heated to 70°C  
30 for 10 minutes, then snap-chilled on ice and spun  
31 briefly to collect the contents to the bottom of the  
32 tube. The following were then added to the tube:

33	1000u/ml RNAGuard	1µl
34	5x first strand buffer	4µl
35	0.1M DTT	2µl
36	10mM dNTPs	1µl

1           200u/ $\mu$ l SuperScript II reverse transcriptase 5 $\mu$ l  
2       The mixture was mixed by pipetting gently and incubated  
3       at 37°C for 1 hour, then placed on ice.

4  
5       (b) Second strand synthesis.

6  
7       The following reagents were added to the first strand  
8       reaction:

9           DEPC-treated water	93 $\mu$ l
10          5x second strand buffer	30 $\mu$ l
11          10mM dNTPs	3 $\mu$ l
12          10u/ $\mu$ l <i>E. coli</i> DNA ligase	1 $\mu$ l
13          10u/ $\mu$ l <i>E. coli</i> DNA polymerase	4 $\mu$ l
14          2u/ $\mu$ l <i>E. coli</i> RNase H	1 $\mu$ l

15       The reaction was vortex mixed and incubated at 16°C for  
16       2 hours. 2 $\mu$ l (10u) of T4 DNA polymerase was added and  
17       incubation continued at 16°C for 5 minutes. The  
18       reaction was placed on ice and 10 $\mu$ l 0.5M EDTA added,  
19       then phenol-chloroform extracted, precipitated and  
20       vacuum dried.

21  
22       (c) Sal I adaptor ligation.

23  
24       The cDNA pellet was resuspended in 25 $\mu$ l DEPC-treated  
25       water, and ligation set up as follows.

26           cDNA	25 $\mu$ l
27           5x T4 DNA ligase buffer	10 $\mu$ l
28           1 $\mu$ g/ $\mu$ l Sal I adapters*	10 $\mu$ l
29           1u/ $\mu$ l T4 DNA ligase	5 $\mu$ l

30       \*Sal I adapters:   TCGACCCACGCGTCCG-3' (SEQ ID No 22)  
31                           GGGTGCCGAGGC-5' (SEQ ID No 23)

32       The ligation was mixed gently and incubated for 16  
33       hours at 16°C, then phenol-chloroform extracted,  
34       precipitated and vacuum dried. The cDNA/adaptor pellet  
35       was resuspended in 41 $\mu$ l of DEPC-treated water and  
36       digested with 60 units of NotI at 37°C for 2 hours,

1 then phenol-chloroform extracted, precipitated and  
2 vacuum dried. The cDNA pellet was re-dissolved in 100 $\mu$ l  
3 TEN buffer (10mM Tris pH 7.5, 0.1mM EDTA, 25mM NaCl)  
4 and size fractionated using a Sephacryl S-500 HR column  
5 to remove unligated adapters and small cDNA fragments  
6 (<400bp) according to the manufacturers instructions.  
7 Fractions were checked by agarose gel electrophoresis  
8 and fractions containing cDNA less than 400 base pairs  
9 discarded, while the remaining fractions were pooled.

10  
11 (d) PCR amplification of oestrogen receptor DNA binding  
12 domain.

13  
14 The oestrogen receptor was PCR amplified from 5 $\mu$ l (150-  
15 250ng) of bone marrow cDNA using 25pmol of each of the  
16 primers pDM12FOR (SEQ ID No 24) (5'-  
17 AAAAGAATTCTGAATGTGTTATTTTAGCTCAGGTCACTCTGACCTGATTATCAAG  
18 ACCCCACTTCACCCCT) and pDM12BAK (SEQ ID No 25) (5'-  
19 AAAAGCGGCCGCAGGGGAGGAGGGTCCATGGAATCTGCCAAGGAG-3') in  
20 two 50 $\mu$ l reactions containing 0.1mM dNTPs, 2.5 units  
21 Taq DNA polymerase, and 1x PCR reaction buffer (10mM  
22 Tris-HCl pH 9.0, 5mM KCl, 0.01% Triton X<sup>®</sup>-100, 1.5mM  
23 MgCl<sub>2</sub>) (Promega Ltd, Southampton, U.K.). The pDM12FOR  
24 primer anneals to the 3'-end of the DNA binding domain  
25 of the oestrogen receptor and incorporates two stop  
26 codons, the 38 base pair consensus oestrogen receptor  
27 HRE sequence, and an EcoRI restriction site. The  
28 pDM12BAK primer anneals to the 5'-end of the DNA  
29 binding domain of the oestrogen receptor and  
30 incorporates the (Gly)<sub>4</sub>Ser linker and the NotI  
31 restriction site.

32  
33 Reactions were overlaid with mineral oil and PCR  
34 carried out on a Techne PHC-3 thermal cycler for 30  
35 cycles of 94°C, 1 minute; 65°C, 1 minute; 72°C, 1  
36 minute. Reaction products were electrophoresed on an

1 agarose gel, excised and products purified from the gel  
2 using a Geneclean II kit according to the manufacturers  
3 instructions (Bio101, La Jolla, California, U.S.A.).

4  
5 (e) Restriction digestion and ligation.

6  
7 The PCR reaction appended NotI and EcoRI restriction  
8 sites, the (Gly)<sub>4</sub>Ser linker, stop codons and the 38 base  
9 pair oestrogen receptor target HRE nucleotide sequence  
10 to the oestrogen receptor DNA binding domain sequence  
11 (see Figure 1). The DNA PCR fragment and the target  
12 pDM6 vector (approximately 500ng) were NotI and EcoRI  
13 digested for 1 hour at 37°C, and DNA purified by  
14 agarose gel electrophoresis and extraction with  
15 Geneclean II kit (Bio101, La Jolla, California,  
16 U.S.A.). The oestrogen receptor DNA binding domain  
17 cassette was ligated into the NotI-EcoRI digested pDM6  
18 vector overnight at 16°C, phenol/chloroform extracted  
19 and precipitated then electroporated into TG1 *E. coli*  
20 (genotype: K12, ( $\Delta$ lac-pro), supE, thi, hsd5/F' traD36,  
21 proA<sup>+</sup>B<sup>+</sup>, LacI<sup>q</sup>, LacZ $\Delta$ 15) and plated onto 2xTY agar  
22 plates supplemented with 1% glucose and 100 $\mu$ g/ml  
23 ampicillin. Colonies were allowed to grow overnight at  
24 37°C. Individual colonies were picked into 5ml 2xTY  
25 supplemented with 1% glucose and 100 $\mu$ g/ml ampicillin  
26 and grown overnight at 37°C. Double stranded phagemid  
27 DNA was isolated with a Wizard SV plasmid DNA isolation  
28 kit and the sequence confirmed with a Prism dyedexoxy  
29 cycle sequencing kit (Perkin-Elmer, Warrington,  
30 Lancashire, U.K.) using M13FOR (SEQ ID No 26) (5'-  
31 GTAAAACGACGGCCAGT) and M13REV (SEQ ID No 27) (5'-  
32 GGATAACAATTTTCACACAGG) oligonucleotides. The pDM12 PDCP  
33 display vector DNA sequence between the HindIII and  
34 EcoRI restriction sites is shown in Figure 1.

35  
36 Example 2. Insertion of a random-primed human

1 lymphocyte cDNA into pDM12 and preparation of a master  
2 PDCP stock.

3

4 Libraries of peptides can be constructed by many  
5 methods known to those skilled in the art. The example  
6 given describes a method for constructing a peptide  
7 library from randomly primed cDNA, prepared from mRNA  
8 isolated from a partially purified cell population.

9

10 mRNA was isolated from approximately  $10^9$  human  
11 peripheral blood lymphocytes using a polyAtract 1000  
12 mRNA isolation kit (Promega, Southampton, UK). The cell  
13 pellet was resuspended in 4ml extraction buffer (4M  
14 guanidine thiocyanate, 25mM sodium citrate pH 7.1, 2%  
15  $\beta$ -mercaptoethanol). 8ml of pre-heated (70°C) dilution  
16 buffer (6xSSC, 10mM Tris pH 7.4, 1mM EDTA, 0.25% SDS,  
17 1%  $\beta$ -mercaptoethanol) was added to the homogenate and  
18 mixed thoroughly by inversion. 10 $\mu$ l of biotinylated  
19 oligo-dT (50 pmol/ $\mu$ l) was added, mixed and the mixture  
20 incubated at 70°C for 5 minutes. The lymphocyte cell  
21 lysate was transferred to 6x 2ml sterile tubes and spun  
22 at 13,000 rpm in a microcentrifuge for ten minutes at  
23 ambient temperature to produce a cleared lysate. During  
24 this centrifugation, streptavidin coated magnetic beads  
25 were resuspended and 6ml transferred to a sterile 50ml  
26 Falcon tube, then placed in the magnetic stand in a  
27 horizontal position until all the beads were captured.  
28 The supernatant was carefully poured off and beads  
29 resuspended in 6ml 0.5xSSC, then the capture repeated.  
30 This wash was repeated 3 times, and beads resuspended  
31 in a final volume of 6ml 0.5xSSC. The cleared lysate  
32 was added to the washed beads, mixed by inversion and  
33 incubated at ambient temperature for 2 minutes, then  
34 beads captured in the magnetic stand in a horizontal  
35 position. The beads were resuspended gently in 2ml  
36 0.5xSSC and transferred to a sterile 2ml screwtop tube,



1 then captured again in the vertical position, and the  
2 wash solution discarded. This wash was repeated twice  
3 more. 1ml of DEPC-treated water was added to the beads  
4 and mixed gently. The beads were again captured and the  
5 eluted mRNA transferred to a sterile tube. 50 $\mu$ l was  
6 electrophoresed to check the quality and quantity of  
7 mRNA, while the remainder was precipitated with 0.1  
8 volumes 3M sodium acetate and three volumes absolute  
9 ethanol at -80°C overnight in 4 aliquots in sterile  
10 1.5ml screwtop tubes.

11

12 Double stranded cDNA was synthesised as described in  
13 Example 1 using 5 $\mu$ g of lymphocyte mRNA as template.  
14 The cDNA was PCR amplified using oligonucleotides  
15 CDNAPCRFOR (SEQ ID No 28) (5'-  
16 AAAGCGGCCGCACTGGCCTGAGAGA), which anneals to the cDNA  
17 synthesis oligonucleotide described in Example 1 which  
18 is present at the 3'-end of all synthesised cDNA  
19 molecules incorporates a NotI restriction site, and an  
20 equimolar mixture of CDNAPCRBAK1, CDNAPCRBAK2 and  
21 CDNAPCRBAK3.

22 CDNAPCRBAK1: (SEQ ID No 29) 5'-

23 AAAAGGCCCAGCCGGCCATGGCCCAGCCCACCACGCGTCCG,

24 CDNAPCRBAK2: (SEQ ID No 30) 5'-

25 AAAAGGCCCAGCCGGCCATGGCCCAGTCCCACCACGCGTCCG,

26 CDNAPCRBAK3: (SEQ ID No 31) 5'-

27 AAAAGGCCCAGCCGGCCATGGCCCAGTACCCACCACGCGTCCG),

28 all three of which anneal to the SalI adaptor sequence  
29 found at the 5'-end of the cDNA and incorporate a SfiI  
30 restriction site at the cDNA 5'-end. Ten PCR reactions  
31 were carried out using 2 $\mu$ l of cDNA (50ng) per reaction  
32 as described in Example 1 using 25 cycles of 94°C, 1  
33 minute; 60°C, 1 minute; 72°C, 2 minutes. The reactions  
34 were pooled and a 20 $\mu$ l aliquot checked by agarose gel  
35 electrophoresis, the remainder was phenol/chloroform  
36 extracted and ethanol precipitated and resuspended in

1 100 $\mu$ l sterile water. 5 $\mu$ g of pDM12 vector DNA and  
2 lymphocyte cDNA PCR product were SfiI-NotI digested  
3 phenol/chloroform extracted and small DNA fragments  
4 removed by size selection on Chromaspin 1000 spin  
5 columns (Clontech, Palo Alto, California, U.S.A.) by  
6 centrifugation at 700g for 2 minutes at room  
7 temperature. Digested pDM12 and lymphocyte cDNA were  
8 ethanol precipitated and ligated together for 16 hours  
9 at 16°C. The ligated DNA was precipitated and  
10 electroporated in to TG1 *E. coli*. Cells were grown in  
11 1ml SOC medium per cuvette used for 1 hour at 37°C, and  
12 plated onto 2xTY agar plates supplemented with 1%  
13 glucose and 100 $\mu$ g/ml ampicillin.  $10^{-4}$ ,  $10^{-5}$  and  $10^{-6}$   
14 dilutions of the electroporated bacteria were also  
15 plated to assess library size. Colonies were allowed to  
16 grow overnight at 30°C.  $2 \times 10^8$  ampicillin resistant  
17 colonies were recovered on the agar plates.  
18 The bacteria were then scraped off the plates into 40ml  
19 2xTY broth supplemented with 20% glycerol, 1% glucose  
20 and 100 $\mu$ g/ml ampicillin. 5ml was added to a 20ml 2xTY  
21 culture broth supplemented with 1% glucose and 100 $\mu$ g/ml  
22 ampicillin and infected with  $10^{11}$  kanamycin resistance  
23 units (kru) M13K07 helper phage at 37°C for 30 minutes  
24 without shaking, then for 30 minutes with shaking at  
25 200rpm. Infected bacteria were transferred to 200ml  
26 2xTY broth supplemented with 25 $\mu$ g/ml kanamycin,  
27 100 $\mu$ g/ml ampicillin, and 20 $\mu$ M IPTG, then incubated  
28 overnight at 37°C, shaking at 200rpm. Bacteria were  
29 pelleted at 4000rpm for 20 minutes in 50ml Falcon  
30 tubes, and 40ml 2.5M NaCl/20% PEG 6000 was added to  
31 200ml of particle supernatant, mixed vigorously and  
32 incubated on ice for 1 hour to precipitate PDCP  
33 particles. Particles were pelleted at 11000rpm for 30  
34 minutes in 250ml Oakridge tubes at 4°C in a Sorvall  
35 RC5B centrifuge, then resuspended in 2ml PBS buffer  
36 after removing all traces of PEG/NaCl with a pipette,

1 then bacterial debris removed by a 5 minute 13500rpm  
2 spin in a microcentrifuge. The supernatant was filtered  
3 through a 0.45µm polysulfone syringe filter and stored  
4 at -20°C.

5  
6 **Example 3. Isolation of human immunoglobulin kappa**  
7 **light chains by repeated rounds of selection against**  
8 **anti-human kappa antibody.**

9  
10 For the first round of library selection a 70x11mm NUNC  
11 Maxisorp Immunotube (Life Technologies, Paisley,  
12 Scotland U.K.) was coated with 2.5ml of 10µg/ml of  
13 anti-human kappa antibody (Seralab, Crawley Down,  
14 Sussex, U.K.) in PBS for 2 hours at 37°C. The tube was  
15 rinsed three times with PBS (fill & empty) and blocked  
16 with 3ml PBS/2% BSA for 2 hours at 37°C and washed as  
17 before.  $4 \times 10^{12}$  a.r.u. of pDM12-lymphocyte cDNA PDCP  
18 stock was added in 2ml 2% BSA/PBS/0.05% Tween 20, and  
19 incubated for 30 minutes on a blood mixer, then for 90  
20 minutes standing at ambient temperature. The tube was  
21 washed ten times with PBS/0.1% Tween 20, then a further  
22 ten times with PBS only. Bound particles were eluted in  
23 1ml of freshly prepared 0.1M triethylamine for 10  
24 minutes at ambient temperature on a blood mixer. Eluted  
25 particles were transferred to 0.5ml 1M Tris pH 7.4,  
26 vortex mixed briefly and transferred to ice.

27  
28 Neutralised particles were added to 10ml log phase TG1  
29 E coli bacteria (optical density:  $OD_{600nm}$  0.3-0.5) and  
30 incubated at 37°C without shaking for 30 minutes, then  
31 with shaking at 200rpm for 30 minutes.  $10^{-3}$ ,  $10^{-4}$  &  $10^{-5}$   
32 dilutions of the infected culture were prepared to  
33 estimate the number of particles recovered, and the  
34 remainder was spun at 4000 rpm for 10 minutes, and the  
35 pellet resuspended in 300µl 2xTY medium by vortex  
36 mixing. Bacteria were plated onto 2xTY agar plates

1 supplemented with 1% glucose and 100 $\mu$ g/ml ampicillin.  
2 Colonies were allowed to grow overnight at 30°C.

3

4 A PDCP stock was prepared from the bacteria recovered  
5 from the first round of selection, as described in  
6 Example 2 from a 100ml overnight culture. 250 $\mu$ l of the  
7 round 1 amplified PDCP stock was then selected against  
8 anti-human kappa antibody as described above with the  
9 tube was washed twelve times with PBS/0.1% Tween 20,  
10 then a further twelve times with PBS only.

11

12 To identify selected clones, eighty-eight individual  
13 clones recovered from the second round of selection  
14 were then tested by ELISA for binding to anti-human  
15 kappa antibody. Individual colonies were picked into  
16 100 $\mu$ l 2xTY supplemented with 100 $\mu$ g/ml ampicillin and 1%  
17 glucose in 96-well plates (Costar) and incubated at  
18 37°C and shaken at 200rpm for 4 hours. 25 $\mu$ l of each  
19 culture was transferred to a fresh 96-well plate,  
20 containing 25 $\mu$ l/well of the same medium plus 10<sup>7</sup> k.r.u.  
21 M13K07 kanamycin resistant helper phage and incubated  
22 at 37°C for 30 minutes without shaking, then incubated  
23 at 37°C and shaken at 200rpm for a further 30 minutes.  
24 160 $\mu$ l of 2xTY supplemented with 100 $\mu$ g/ml ampicillin,  
25 25 $\mu$ g/ml kanamycin, and 20 $\mu$ M IPTG was added to each well  
26 and particle amplification continued for 16 hours at  
27 37°C while shaking at 200rpm. Bacterial cultures were  
28 spun in microtitre plate carriers at 2000g for 10  
29 minutes at 4°C in a benchtop centrifuge to pellet  
30 bacteria and culture supernatant used for ELISA.

31

32 A Dynatech Immulon 4 ELISA plate was coated with  
33 200ng/well anti-human kappa antibody in 100 $\mu$ l /well PBS  
34 for one hour at 37°C. The plate was washed 2x200 $\mu$ l/well  
35 PBS and blocked for 1 hour at 37°C with 200 $\mu$ l/well 2%  
36 BSA/PBS and then washed 2x200 $\mu$ l/well PBS. 50 $\mu$ l PDCP

1 culture supernatant was added to each well containing  
2 50 $\mu$ l/well 4% BSA/PBS/0.1%Tween 20, and allowed to bind  
3 for 1 hour at ambient temperature. The plate was washed  
4 three times with 200 $\mu$ l/well PBS/0.1% Tween 20, then  
5 three times with 200 $\mu$ l/well PBS. Bound PDCPs were  
6 detected with 100 $\mu$ l/well, 1:5000 diluted anti-M13-HRP  
7 conjugate (Pharmacia) in 2% BSA/PBS/0.05% Tween 20 for  
8 1 hour at ambient temperature and the plate washed six  
9 times as above. The plate was developed for 5 minutes  
10 at ambient temperature with 100 $\mu$ l/well freshly prepared  
11 TMB (3,3',5,5'-Tetramethylbenzidine) substrate buffer  
12 (0.005% H<sub>2</sub>O<sub>2</sub>, 0.1mg/ml TMB in 24mM citric acid/52mM  
13 sodium phosphate buffer pH 5.2). The reaction was  
14 stopped with 100 $\mu$ l/well 12.5% H<sub>2</sub>SO<sub>4</sub> and read at 450nm.  
15 (ELISA data for binding clones is shown in Figure 2).  
16

17 These clones were then sequenced with M13REV primer  
18 (SEQ ID No 27) as in Example 1. The sequence of two of  
19 the clones isolated is shown in Figure 3 (see SEQ ID  
20 Nos 7 to 10).  
21

#### 22 Example 4. Construction of the pDM14 N-terminal display 23 vector 24

25 It would be useful to design vectors that contain a  
26 second DBD binding sequence, such as a second oestrogen  
27 receptor HRE sequence, thus allowing the display of  
28 increased numbers of peptides per PDCP. Peale et al.  
29 (1988, Proc. Natl. Acad. Sci. USA 85: 1038-1042)  
30 describe a number of oestrogen receptor HRE sequences.  
31 These sequences were used to define an HRE sequence,  
32 which differs from that cloned in pDM12, which we used  
33 to create a second N-terminal display vector (pDM14).  
34 The oligonucleotide: 5'-AAAAGAATTCGAGGTTACATTAAGTTTGT  
35 CCGGTCAGACTGACCCAAGTCGACCTGAATGTGTTATTTTAG-3' (SEQ ID  
36 No 32) was synthesised and used to mutagenise pDM12 by

1 PCR with pDM12BAK oligonucleotide as described in  
2 Example 1 using 100ng pDM12 vector DNA as template. The  
3 resulting DNA fragment, which contained the oestrogen  
4 receptor DBD and two HRE sequences separated by a SalI  
5 restriction enzyme site, was NotI-EcoRI restriction  
6 enzyme digested and cloned into NotI-EcoRI digested  
7 pDM12 vector DNA as described in Example 1 to create  
8 pDM14. The sequence of pDM14 between the HindIII and  
9 EcoRI restriction enzyme sites was checked by DNA  
10 sequencing. The final vector sequence between these two  
11 sites is shown in Figure 4 (see SEQ ID Nos 11 and 12).

12  
13 **Example 5. Construction of the pDM16 C-terminal display**  
14 **vector**

15  
16 In order to demonstrate the display of peptides fused  
17 to the C-terminus of a DBD on a PDCP a suitable vector,  
18 pDM16, was created.

19  
20 In pDM16 the pelB leader DNA sequence is fused directly  
21 to the oestrogen receptor DBD sequence removing the  
22 multiple cloning sites and the Gly<sub>4</sub>Ser linker DNA  
23 sequence found in pDM12 and pDM14, which are appended  
24 to the C-terminal end of the DBD sequence upstream of  
25 the HRE DNA sequence.

26  
27 To create this vector two separate PCR reactions were  
28 carried out on a Techne Progene thermal cycler for 30  
29 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1  
30 minute. Reaction products were electrophoresed on an  
31 agarose gel, excised and products purified from the gel  
32 using a Mermaid or Geneclean II kit, respectively,  
33 according to the manufacturers instructions (Bio101, La  
34 Jolla, California, U.S.A.).

35  
36 In the first, the 5'-untranslated region and pelB

1 leader DNA sequence was amplified from 100ng of pDM12  
2 vector DNA using 50pmol of each of the oligonucleotides  
3 pelBFOR (SEQ ID No 33) (5'-CCTTGGCAGATTCCATCT  
4 CGGCCATTGCCGGC-3') and M13REV (SEQ ID NO 27) (see  
5 above) in a 100 $\mu$ l reaction containing 0.1mM dNTPs, 2.5  
6 units Taqplus DNA polymerase, and 1x High Salt PCR  
7 reaction buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM  
8 MgCl<sub>2</sub>) (Stratagene Ltd, Cambridge, U.K.).

9

10 In the second, the 3'-end of the pelB leader sequence  
11 and the oestrogen receptor DBD was amplified from 100ng  
12 of pDM12 vector DNA using 50pmol of each of the  
13 oligonucleotides pelBBAK (SEQ ID No 34) (5'-CCGGCAA  
14 TGGCCGAGATGGAATCTGCCAAGG-3') and pDM16FOR (SEQ ID No  
15 35) (5'-TTTTGTCTGACTCAATCAGTTATGCGGCCGCCAGCTGCAGG  
16 AGGGCCGGCTGGGCGGACCCTCCTCCCCAGACCCCACTTCACCCC-3') in a  
17 100 $\mu$ l reaction containing 0.1mM dNTPs, 2.5 units  
18 Taqplus DNA polymerase, and 1x High Salt PCR reaction  
19 buffer (Stratagene Ltd, Cambridge, U.K.). Following gel  
20 purification both products were mixed together and a  
21 final round of PCR amplification carried out to link  
22 the two products together as described above, in a  
23 100 $\mu$ l reaction containing 0.1mM dNTPs, 2.5 units Taq  
24 DNA polymerase, and 1x PCR reaction buffer (10mM Tris-  
25 HCl pH 9.0, 5mM KCl, 0.01% Triton X-100, 1.5mM MgCl<sub>2</sub>)  
26 (Promega Ltd, Southampton, U.K.).

27

28 The resulting DNA fragment, was HindIII-SalI  
29 restriction enzyme digested and cloned into HindIII-  
30 SalI digested pDM14 vector DNA as described in Example  
31 1 to create pDM16. The sequence of pDM16 between the  
32 HindIII and EcoRI restriction enzyme sites was checked  
33 by DNA sequencing. The final vector sequence between  
34 these two sites is shown in Figure 5 (see SEQ ID Nos 13  
35 and 14).

36

1     **Example 6. Display of the C-terminal fragment of human**  
2     **N-cadherin on the surface of a PDCP**

3  
4     cDNA libraries of peptides can be constructed by many  
5     methods known to those skilled in the art. One commonly  
6     used method for constructing a peptide library uses  
7     oligo dT primed cDNA, prepared from polyA+ mRNA. In  
8     this method the first-strand synthesis is carried out  
9     using an oligonucleotide which anneals to the 3'-end  
10    polyA tail of the mRNA composed of  $T_n$  (where n is  
11    normally between 10 and 20 bases) and a restriction  
12    enzyme site such as NotI to facilitate cloning of cDNA.  
13    The cDNA cloned by this method is normally composed of  
14    the polyA tail, the 3'- end untranslated region and the  
15    C-terminal coding region of the protein. As an example  
16    of the C-terminal display of peptides on a PDCP, a  
17    human cDNA isolated from a library constructed by the  
18    above method was chosen.

19  
20    The protein N-cadherin is a cell surface molecule  
21    involved in cell-cell adhesion. The C-terminal  
22    cytoplasmic domain of the human protein (Genbank  
23    database accession number: M34064) is recognised by a  
24    commercially available monoclonal antibody which was  
25    raised against the C-terminal 23 amino acids of chicken  
26    N-cadherin (SIGMA catalogue number: C-1821). The 1.4kb  
27    human cDNA fragment encoding the C-terminal 99 amino  
28    acids, 3'- untranslated region and polyA tail (NotI  
29    site present at the 3'-end of the polyA tail) was  
30    amplified from approximately 20ng pDM7-NCAD#C with  
31    25pmol of each oligonucleotide M13FOR (SEQ ID No 26)  
32    and CDNPCRBAK1 (SEQ ID No 29) (see above) in a 50 $\mu$ l  
33    reaction containing 0.1mM dNTPs, 2.5 units Taqplus DNA  
34    polymerase, and 1x High Salt PCR reaction buffer (20mM  
35    Tris-HCl pH 9.2, 60mM KCl, 2mM MgCl<sub>2</sub>) (Stratagene Ltd,  
36    Cambridge, U.K.) on a Techne Progene thermal cycler for



1 30 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1  
2 minute. Following gel purification and digestion with  
3 SfiI and NotI restriction enzymes, the PCR product was  
4 cloned into pDM16 using an analogous protocol as  
5 described in Example 1.

6  
7 Clones containing inserts were identified by ELISA of  
8 96 individual PDCP cultures prepared as described in  
9 Example 3. A Dynatech Immulon 4 ELISA plate was coated  
10 with 1:250 diluted anti-pan cadherin monoclonal  
11 antibody in 100µl /well PBS overnight at 4°C. The plate  
12 was washed 3x200µl/well PBS and blocked for 1 hour at  
13 37°C with 200µl/well 2% Marvel non-fat milk powder/PBS  
14 and then washed 2x200µl/well PBS. 50µl PDCP culture  
15 supernatant was added to each well containing 50µl/well  
16 4% Marvel/PBS, and allowed to bind for 1 hour at  
17 ambient temperature. The plate was washed three times  
18 with 200µl/well PBS/0.1% Tween 20, then three times  
19 with 200µl/well PBS. Bound PDCPs were detected with  
20 100µl/well, 1:5000 diluted anti-M13-HRP conjugate  
21 (Pharmacia) in 2% Marvel/PBS for 1 hour at ambient  
22 temperature and the plate washed six times as above.  
23 The plate was developed for 15 minutes at ambient  
24 temperature with 100µl/well freshly prepared TMB  
25 (3,3',5,5'-Tetramethylbenzidine) substrate buffer  
26 (0.005% H<sub>2</sub>O<sub>2</sub>, 0.1mg/ml TMB in 24mM citric acid/52mM  
27 sodium phosphate buffer pH 5.2). The reaction was  
28 stopped with 100µl/well 12.5% H<sub>2</sub>SO<sub>4</sub> and read at 450nm.  
29 The nucleotide sequence of an ELISA positive clone  
30 insert and DBD junction was checked by DNA sequencing  
31 using oligonucleotides M13FOR (SEQ ID No 26) (see  
32 Example 1) and ORSEQBAK (SEQ ID No 36) (5'-  
33 TGTGAAACACAAGCGCCAG-3').

34  
35 A fifty-fold concentrated stock of C-terminal N-  
36 cadherin PDCP particles was prepared by growing the un-

1 infected TG1 clone in 1ml 2xTY culture broth  
2 supplemented with 1% glucose and 100µg/ml ampicillin  
3 for five hours at 37°C, shaking at 200rpm and infecting  
4 with 10<sup>8</sup> kanamycin resistance units (kru) M13K07 helper  
5 phage at 37°C for 30 minutes without shaking, then for  
6 30 minutes with shaking at 200rpm. Infected bacteria  
7 were transferred to 20ml 2xTY broth supplemented with  
8 25µg/ml kanamycin, 100µg/ml ampicillin, and 20µM IPTG,  
9 then incubated overnight at 30°C, shaking at 200rpm.  
10 Bacteria were pelleted at 4000rpm for 20 minutes in  
11 50ml Falcon tubes, and 4ml 2.5M NaCl/20% PEG 6000 was  
12 added to 20ml of PDCP supernatant, mixed vigorously and  
13 incubated on ice for 1 hour to precipitate particles.

14  
15 The particles were pelleted at 11000rpm for 30 minutes  
16 in 50ml Oakridge tubes at 4°C in a Sorvall RC5B  
17 centrifuge, then resuspended in PBS buffer after  
18 removing all traces of PEG/NaCl with a pipette, then  
19 bacterial debris removed by a 5 minute 13500rpm spin in  
20 a microcentrifuge. The supernatant was filtered through  
21 a 0.45µm polysulfone syringe filter. The concentrated  
22 stock was two-fold serially diluted and used in ELISA  
23 against plates coated with anti-pan-cadherin antibody  
24 as described above (see Figure 6).

25  
26 This example demonstrates the principle of C-terminal  
27 display using PDCPs, that C-terminal DBD-peptide fusion  
28 PDCPs can be made which can be detected in ELISA, and  
29 the possibility that oligo dT primed cDNA libraries may  
30 be displayed using this method.

31  
32 Example 7. Display of *in vivo* biotinylated C-terminal  
33 domain of human propionyl CoA carboxylase on the  
34 surface of a PDCP

35

36 Example 6 shows that the C-terminal domain of human N-

1 cadherin can be expressed on the surface of a PDCP as a  
2 C-terminal fusion with the DBD. Here it is shown that  
3 the C-terminal domain of another human protein  
4 propionyl CoA carboxylase alpha chain (Genbank  
5 accession number: X14608) can similarly be displayed,  
6 suggesting that this methodology may be general.

7  
8 The alpha sub-unit of propionyl CoA carboxylase alpha  
9 chain (PCC) contains 703 amino acids and is normally  
10 biotinylated at position 669. It is demonstrated that  
11 the PCC peptide displayed on the PDCP is biotinylated,  
12 as has been shown to occur when the protein is  
13 expressed in bacterial cells (Leon-Del-Rio & Gravel;  
14 1994, J. Biol. Chem. 37, 22964-22968).

15  
16 The 0.8kb human cDNA fragment of PCC alpha encoding the  
17 C-terminal 95 amino acids, 3'- untranslated region and  
18 polyA tail (NotI site present at the 3'-end of the  
19 polyA tail) was amplified and cloned into pDM16 from  
20 approximately 20ng pDM7-PCC#C with 25pmol of each  
21 oligonucleotide M13FOR (SEQ ID No 26) and CDNPCRBAK1  
22 (SEQ ID No 29) as described in Example 6.

23  
24 Clones containing inserts were identified by ELISA as  
25 described in Example 6, except that streptavidin was  
26 coated onto the ELISA plate at 250ng/well, in place of  
27 the anti-cadherin antibody. The nucleotide sequence of  
28 an ELISA positive clone insert and DBD junction was  
29 checked by DNA sequencing using oligonucleotides M13FOR  
30 (SEQ ID No 26) and ORSEQBAK (SEQ ID No 36) (see above).  
31 A fifty-fold concentrated stock of C-terminal PCC PDCP  
32 particles was prepared and tested in ELISA against  
33 streptavidin as described in Example 6 (see Figure 7).

34  
35 This example shows not only that the peptide can be  
36 displayed as a C-terminal fusion on a PDCP, but also

1 that in vivo modified peptides can be displayed.

2

3 **Example 8. Construction of a human scFv PDCP display**  
4 **library**

5

6 This example describes the generation of a human  
7 antibody library of scFvs made from an un-immunised  
8 human. The overall strategy for the PCR assembly of  
9 scFv fragments is similar to that employed by Marks, J.  
10 D. et al. 1991, J. Mol. Biol. 222: 581-597. The  
11 antibody gene oligonucleotides used to construct the  
12 library are derived from the Marke et al., paper and  
13 from sequence data extracted from the Kabat database  
14 (Kabat, E. A. et al., Sequences of Proteins of  
15 Immunological Interest. 4<sup>th</sup> edition. U.S. Department of  
16 Health and Human Services. 1987). The three linker  
17 oligonucleotides are described by Zhou et al. (1994,  
18 Nucleic Acids Res., 22: 888-889), all oligonucleotides  
19 used are detailed in Table 1.

20

21 First, mRNA was isolated from peripheral blood  
22 lymphocytes and cDNA prepared for four repertoires of  
23 antibody genes IgD, IgM, Ig $\kappa$  and Ig $\lambda$ , using four  
24 separate cDNA synthesis primers. VH genes were  
25 amplified from IgD and IgM primed cDNA, and VL genes  
26 were amplified from Ig $\kappa$  and Ig $\lambda$  primed cDNA. A portion  
27 of each set of amplified heavy chain or light chain DNA  
28 was then spliced with a separate piece of linker DNA  
29 encoding the 15 amino acids (Gly, Ser), (Huston, J. S.  
30 et al. 1989, Gene, 77: 61). The 3'-end of the VH PCR  
31 products and the 5'-end of the VL PCR products overlap  
32 the linker sequence as a result of incorporating linker  
33 sequence in the JH, V $\kappa$  and V $\lambda$  family primer sets (Table  
34 1). Each VH-linker or linker-VL DNA product was then  
35 spliced with either VH or VL DNA to produce the primary  
36 scFv product in a VH-linker-VL configuration. This scFv

1 product was then amplified and cloned into pDM12 as a  
2 SfiI-NotI fragment, electroporated into TG1 and a  
3 concentrated PDCP stock prepared.

4  
5 **mRNA isolation and cDNA synthesis.**

6 Human lymphocyte mRNA was purified as described in  
7 Example 2. Separate cDNA reactions were performed with  
8 IGDcDNAFOR (SEQ ID No 37), IGMcDNAFOR (SEQ ID No 38),  
9 IGKcDNAFOR (SEQ ID No 39) and IGLcDNAFOR (SEQ ID No 40)  
10 oligonucleotides. 50pmol of each primer was added to  
11 approximately 5µg of mRNA in 20µl of nuclease free  
12 water and heated to 70°C for 5 minutes and cooled  
13 rapidly on ice, then made up to a final reaction volume  
14 of 100µl containing 50mM Tris pH 8.3, 75mM KCl, 3mM  
15 MgCl<sub>2</sub>, 10mM DTT, 0.5mM dNTPs, and 2000 units of  
16 Superscript II reverse transcriptase (Life  
17 Technologies, Paisley, Scotland, U.K.). The reactions  
18 were incubated at 37°C for two hours, then heated to  
19 95°C for 5 minutes.

20  
21 **Primary PCRs.**

22 For the primary PCR amplifications separate  
23 amplifications were set up for each family specific  
24 primer with either an equimolar mixture of the JHFOR  
25 primer set (SEQ ID Nos 41 to 44) for IgM and IgD cDNA,  
26 or with SCFVκFOR (SEQ ID No 51) or SCFVλFOR (SEQ ID No  
27 52) for IgK or Igλ cDNA respectively e.g. VH1BAK and  
28 JHFOR set; Vκ2BAK (SEQ ID No 54) and SCFVκFOR (SEQ ID  
29 No 51); Vλ3aBAK (SEQ ID No 66) and SCFVλFOR (SEQ ID No  
30 52) etc. Thus, for IgM, IgD and Igκ cDNA six separate  
31 reactions were set up, and seven for Igλ cDNA. A 50µl  
32 reaction mixture was prepared containing 2µl cDNA,  
33 25pmol of the appropriate FOR and BAK primers, 0.1mM  
34 dNTPs, 2.5 units Taqplus DNA polymerase, and 1x High  
35 Salt PCR reaction buffer (20mM Tris-HCl pH 9.2, 60mM  
36 KCl, 2mM MgCl<sub>2</sub>) (Stratagene Ltd, Cambridge, U.K.).

1 1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM KCl,  
2 0.01% Triton X<sup>®</sup>-100, 1.5mM MgCl<sub>2</sub>) (Promega Ltd,  
3 Southampton, U.K.). Reactions were amplified on a  
4 Techne Progene thermal cycler for 30 cycles of 94°C, 1  
5 minute; 60°C, 1 minute; 72°C, 2 minutes, followed by 10  
6 minutes at 72°C. Reaction products were electrophoresed  
7 on an agarose gel, excised and products purified from  
8 the gel using a Geneclean II kit according to the  
9 manufacturers instructions (Bio101, La Jolla,  
10 California, U.S.A.) and adjusted to 20ng/μl.

11  
12 **Final linking and reamplification.**

13 To prepare the final scFv DNA products, five 100μl  
14 reactions were performed for VH-LINKER plus VL DNA,  
15 and, five 100μl reactions were performed for VH plus  
16 LINKER-VL DNA for each of the four final repertoires  
17 (IgM VH-VK, VH-Vλ; IgD VH-VK, VH-Vλ) as described in  
18 step (d) above using 20ng of each component DNA as  
19 template. Reaction products were electrophoresed on an  
20 agarose gel, excised and products purified from the gel  
21 using a Geneclean II kit according to the manufacturers  
22 instructions (Bio101, La Jolla, California, U.S.A.) and  
23 adjusted to 20ng/μl. Each of the four repertoires was  
24 then re-amplified in a 100μl reaction volume containing  
25 2ng of each linked product, with 50pmol VHBK1-6 (SEQ  
26 ID Nos 53 to 58) and either the JKFOR (SEQ ID Nos 66 to  
27 70) or JλFOR (SEQ ID Nos 71 to 73) primer sets, in the  
28 presence of 0.1mM dNTPs, 2.5 units Taq DNA polymerase,  
29 and 1x PCR reaction buffer (10mM Tris-HCl pH 9.0, 5mM  
30 KCl, 0.01% Triton X<sup>®</sup>-100, 1.5mM MgCl<sub>2</sub>) (Promega Ltd,  
31 Southampton, U.K.). Thirty reactions were performed per  
32 repertoire to generate enough DNA for cloning.  
33 Reactions were amplified on a Techne Progene thermal  
34 cycler for 25 cycles of 94°C, 1 minute; 65°C, 1 minute;  
35 72°C, 2 minutes, followed by 10 minutes at 72°C.  
36 Reaction products were phenol-chloroform extracted,

1 Reactions were amplified on a Techne Progene thermal  
2 cyclor for 30 cycles of 94°C, 1 minute; 60°C, 1 minute;  
3 72°C, 2 minutes, followed by 10 minutes at 72°C. Fifty  
4 microlitres of all 25 reaction products were  
5 electrophoresed on an agarose gel, excised and products  
6 purified from the gel using a Geneclean II kit  
7 according to the manufacturers instructions (Bio101, La  
8 Jolla, California, U.S.A.). All sets of IgD, IgM, IgK  
9 or Igλ reaction products were pooled to produce VH or  
10 VL DNA sets for each of the four repertoires. These  
11 were then adjusted to approximately 20ng/μl.

12  
13 **Preparation of linker.**

14 Linker product was prepared from eight 100μl reactions  
15 containing 5ng LINKAMP3T (SEQ ID No 76) template  
16 oligonucleotide, 50pmol of LINKAMP3 (SEQ ID No 74) and  
17 LINKAMP5 (SEQ ID No 75) primers, 0.1mM dNTPs, 2.5 units  
18 Taqplus DNA polymerase, and 1x High Salt PCR reaction  
19 buffer (20mM Tris-HCl pH 9.2, 60mM KCl, 2mM MgCl<sub>2</sub>)  
20 (Stratagene Ltd, Cambridge, U.K.). Reactions were  
21 amplified on a Techne Progene thermal cyclor for 30  
22 cycles of 94°C, 1 minute; 60°C, 1 minute; 72°C, 1  
23 minute, followed by 10 minutes at 72°C. All reaction  
24 product was electrophoresed on a 2% low melting point  
25 agarose gel, excised and products purified from the gel  
26 using a Mermaid kit according to the manufacturers  
27 instructions (Bio101, La Jolla, California, U.S.A.) and  
28 adjusted to 5ng/μl.

29  
30 **First stage linking.**

31 Four linking reactions were prepared for each  
32 repertoire using 20ng of VH or VL DNA with 5ng of  
33 Linker DNA in 100μl reactions containing (for IgM or  
34 IgD VH) 50pmol of LINKAMPFOR and VH1-6BAK set, or,  
35 50pmol LINKAMPBAK and either SCFVκFOR (Igκ) or SCFVλFOR  
36 (Igλ), 0.1mM dNTPs, 2.5 units Taq DNA polymerase, and

1 ethanol precipitated, vacuum dried and re-suspended in  
2 80 $\mu$ l nuclease free water.

3

#### 4 Cloning into pDM12.

5 Each of the four repertoires was SfiI-NotI digested,  
6 and electrophoresed on an agarose gel, excised and  
7 products purified from the gel using a Geneclean II kit  
8 according to the manufacturers instructions (Bio101, La  
9 Jolla, California, U.S.A.). Each of the four  
10 repertoires was ligated overnight at 16°C in 140 $\mu$ l with  
11 10 $\mu$ g of SfiI-NotI cut pDM12 prepared as in Example 2,  
12 and 12 units of T4 DNA ligase (Life Technologies,  
13 Paisley, Scotland, U.K.). After incubation the  
14 ligations were adjusted to 200 $\mu$ l with nuclease free  
15 water, and DNA precipitated with 1 $\mu$ l 20mg/ml glycogen,  
16 100 $\mu$ l 7.5M ammonium acetate and 900 $\mu$ l ice-cold (-20°C)  
17 absolute ethanol, vortex mixed and spun at 13,000rpm  
18 for 20 minutes in a microfuge to pellet DNA. The  
19 pellets were washed with 500 $\mu$ l ice-cold 70% ethanol by  
20 centrifugation at 13,000rpm for 2 minutes, then vacuum  
21 dried and re-suspended in 10 $\mu$ l DEPC-treated water. 1 $\mu$ l  
22 aliquots of each repertoire was electroporated into  
23 80 $\mu$ l *E. coli* (TG1). Cells were grown in 1ml SOC medium  
24 per cuvette used for 1 hour at 37°C, and plated onto  
25 2xTY agar plates supplemented with 1% glucose and  
26 100 $\mu$ g/ml ampicillin. 10<sup>-4</sup>, 10<sup>-5</sup> and 10<sup>-6</sup> dilutions of the  
27 electroporated bacteria were also plated to assess  
28 library size. Colonies were allowed to grow overnight  
29 at 30°C. Cloning into SfiI-NotI digested pDM12 yielded  
30 an IgM- $\kappa/\lambda$  repertoire of 1.16x10<sup>9</sup> clones, and an IgD- $\kappa/\lambda$   
31 repertoire of 1.21x10<sup>9</sup> clones.

32

#### 33 Preparation of PDCP stock.

34 Separate PDCP stocks were prepared for each repertoire  
35 library. The bacteria were then scraped off the plates  
36 into 30ml 2xTY broth supplemented with 20% glycerol, 1%



1 glucose and 100 $\mu$ g/ml ampicillin. 3ml was added to a  
2 50ml 2xTY culture broth supplemented with 1% glucose  
3 and 100 $\mu$ g/ml ampicillin and infected with 10<sup>11</sup> kanamycin  
4 resistance units (kru) M13K07 helper phage at 37°C for  
5 30 minutes without shaking, then for 30 minutes with  
6 shaking at 200rpm. Infected bacteria were transferred  
7 to 500ml 2xTY broth supplemented with 25 $\mu$ g/ml  
8 kanamycin, 100 $\mu$ g/ml ampicillin, and 20 $\mu$ M IPTG, then  
9 incubated overnight at 30°C, shaking at 200rpm.  
10 Bacteria were pelleted at 4000rpm for 20 minutes in  
11 50ml Falcon tubes, and 80ml 2.5M NaCl/20% PEG 6000 was  
12 added to 400ml of particle supernatant, mixed  
13 vigorously and incubated on ice for 1 hour to  
14 precipitate PDCP particles. Particles were pelleted at  
15 11000rpm for 30 minutes in 250ml Oakridge tubes at 4°C  
16 in a Sorvall RC5B centrifuge, then resuspended in 40ml  
17 water and 8ml 2.5M NaCl/20% PEG 6000 added to  
18 reprecipitate particles, then incubated on ice for 20  
19 minutes. Particles were again pelleted at 11000rpm for  
20 30 minutes in 50ml Oakridge tubes at 4°C in a Sorvall  
21 RC5B centrifuge, then resuspended in 5ml PBS buffer,  
22 after removing all traces of PEG/NaCl with a pipette.  
23 Bacterial debris was removed by a 5 minute 13500rpm  
24 spin in a microcentrifuge. The supernatant was filtered  
25 through a 0.45 $\mu$ m polysulfone syringe filter, adjusted  
26 to 20% glycerol and stored at -70°C.

27  
28 **Example 9. Isolation of binding activity from a N-**  
29 **terminal display PDCP library of human scFvs**

30  
31 The ability to select binding activities to a target of  
32 interest from a human antibody library is important due  
33 to the possibility of generating therapeutic human  
34 antibodies. In addition, such libraries allow the  
35 isolation of antibodies to targets which cannot be used  
36 for traditional methods of antibody generation due to

1 ethanol precipitated, vacuum dried and re-suspended in  
2 80 $\mu$ l nuclease free water.

3

#### 4 Cloning into pDM12.

5 Each of the four repertoires was SfiI-NotI digested,  
6 and electrophoresed on an agarose gel, excised and  
7 products purified from the gel using a Geneclean II kit  
8 according to the manufacturers instructions (Bio101, La  
9 Jolla, California, U.S.A.). Each of the four  
10 repertoires was ligated overnight at 16°C in 140 $\mu$ l with  
11 10 $\mu$ g of SfiI-NotI cut pDM12 prepared as in Example 2,  
12 and 12 units of T4 DNA ligase (Life Technologies,  
13 Paisley, Scotland, U.K.). After incubation the  
14 ligations were adjusted to 200 $\mu$ l with nuclease free  
15 water, and DNA precipitated with 1 $\mu$ l 20mg/ml glycogen,  
16 100 $\mu$ l 7.5M ammonium acetate and 900 $\mu$ l ice-cold (-20°C)  
17 absolute ethanol, vortex mixed and spun at 13,000rpm  
18 for 20 minutes in a microfuge to pellet DNA. The  
19 pellets were washed with 500 $\mu$ l ice-cold 70% ethanol by  
20 centrifugation at 13,000rpm for 2 minutes, then vacuum  
21 dried and re-suspended in 10 $\mu$ l DEPC-treated water. 1 $\mu$ l  
22 aliquots of each repertoire was electroporated into  
23 80 $\mu$ l *E. coli* (TG1). Cells were grown in 1ml SOC medium  
24 per cuvette used for 1 hour at 37°C, and plated onto  
25 2xTY agar plates supplemented with 1% glucose and  
26 100 $\mu$ g/ml ampicillin. 10<sup>-4</sup>, 10<sup>-5</sup> and 10<sup>-6</sup> dilutions of the  
27 electroporated bacteria were also plated to assess  
28 library size. Colonies were allowed to grow overnight  
29 at 30°C. Cloning into SfiI-NotI digested pDM12 yielded  
30 an IgM- $\kappa/\lambda$  repertoire of 1.16x10<sup>9</sup> clones, and an IgD- $\kappa/\lambda$   
31 repertoire of 1.21x10<sup>9</sup> clones.

32

#### 33 Preparation of PDCP stock.

34 Separate PDCP stocks were prepared for each repertoire  
35 library. The bacteria were then scraped off the plates  
36 into 30ml 2xTY broth supplemented with 20% glycerol, 1%

1 glucose and 100 $\mu$ g/ml ampicillin. 3ml was added to a  
2 50ml 2xTY culture broth supplemented with 1% glucose  
3 and 100 $\mu$ g/ml ampicillin and infected with 10<sup>11</sup> kanamycin  
4 resistance units (kru) M13K07 helper phage at 37°C for  
5 30 minutes without shaking, then for 30 minutes with  
6 shaking at 200rpm. Infected bacteria were transferred  
7 to 500ml 2xTY broth supplemented with 25 $\mu$ g/ml  
8 kanamycin, 100 $\mu$ g/ml ampicillin, and 20 $\mu$ M IPTG, then  
9 incubated overnight at 30°C, shaking at 200rpm.  
10 Bacteria were pelleted at 4000rpm for 20 minutes in  
11 50ml Falcon tubes, and 80ml 2.5M NaCl/20% PEG 6000 was  
12 added to 400ml of particle supernatant, mixed  
13 vigorously and incubated on ice for 1 hour to  
14 precipitate PDCP particles. Particles were pelleted at  
15 11000rpm for 30 minutes in 250ml Oakridge tubes at 4°C  
16 in a Sorvall RC5B centrifuge, then resuspended in 40ml  
17 water and 8ml 2.5M NaCl/20% PEG 6000 added to  
18 reprecipitate particles, then incubated on ice for 20  
19 minutes. Particles were again pelleted at 11000rpm for  
20 30 minutes in 50ml Oakridge tubes at 4°C in a Sorvall  
21 RC5B centrifuge, then resuspended in 5ml PBS buffer,  
22 after removing all traces of PEG/NaCl with a pipette.  
23 Bacterial debris was removed by a 5 minute 13500rpm  
24 spin in a microcentrifuge. The supernatant was filtered  
25 through a 0.45 $\mu$ m polysulfone syringe filter, adjusted  
26 to 20% glycerol and stored at -70°C.

27  
28 **Example 9. Isolation of binding activity from a N-**  
29 **terminal display PDCP library of human scFvs**

30  
31 The ability to select binding activities to a target of  
32 interest from a human antibody library is important due  
33 to the possibility of generating therapeutic human  
34 antibodies. In addition, such libraries allow the  
35 isolation of antibodies to targets which cannot be used  
36 for traditional methods of antibody generation due to

1 toxicity, low immunogenicity or ethical considerations.  
2 In this example we demonstrate the isolation of  
3 specific binding activities against a peptide antigen  
4 from a PDCP library of scFvs from an un-immunised  
5 human.

6  
7 The generation of the library, used for the isolation  
8 of binding activities in this example, is described in  
9 Example 8.

10  
11 Substance P is an eleven amino acid neuropeptide  
12 involved in inflammatory and pain responses *in vivo*. It  
13 has also been implicated in a variety of disorders such  
14 as psoriasis and asthma amongst others (Misery, L.  
15 1997, Br. J. Dermatol., 137: 843-850; Maggi, C. A.  
16 1997, Regul. Pept. 70: 75-90; Choi, D. C. & Kwon, O.J.,  
17 1998, Curr. Opin. Pulm. Med., 4: 16-24). Human  
18 antibodies which neutralise this peptide may therefore  
19 have some therapeutic potential. As this peptide is too  
20 small to coat efficiently on a tube, as described in  
21 Example 3, selection of binding activities was  
22 performed in-solution, using N-terminal biotinylated  
23 substance P and capturing bound PDCP particles on  
24 streptavidin-coated magnetic beads.

25  
26 **Enrichment for substance P binding PDCP particles.**  
27 An aliquot of approximately  $10^{13}$  a.r.u. IgM and IgD scFv  
28 library stock was mixed with  $1\mu\text{g}$  biotinylated substance  
29 P in  $800\mu\text{l}$  4% BSA/0.1% Tween 20/PBS, and allowed to  
30 bind for two hours at ambient temperature. Bound PDCPs  
31 were then captured onto  $1\text{ml}$  of BSA blocked streptavidin  
32 coated magnetic beads for 10 minutes at ambient  
33 temperature. The beads were captured to the side of the  
34 tube with a magnet (Promega), and unbound material  
35 discarded. The beads were washed eight times with  $1\text{ml}$   
36 PBS/0.1% Tween 20/  $10\mu\text{g/ml}$  streptavidin, then two times

1 with 1ml of PBS by magnetic capture and removal of wash  
2 buffer. After the final wash bound PDCPs were eluted  
3 with 1ml of freshly prepared 0.1M triethylamine for 10  
4 minutes, the beads were captured, and eluted particles  
5 transferred to 0.5ml 1M Tris-HCl pH 7.4. Neutralised  
6 particles were added to 10ml log phase TG1 *E. coli*  
7 bacteria and incubated at 37°C without shaking for 30  
8 minutes, then with shaking at 200rpm for 30 minutes.  
9  $10^{-3}$ ,  $10^{-4}$  &  $10^{-5}$  dilutions of the infected culture were  
10 prepared to estimate the number of particles recovered,  
11 and the remainder was spun at 4000 rpm for 10 minutes,  
12 and the pellet resuspended in 300 $\mu$ l 2xTY medium by  
13 vortex mixing. Bacteria were plated onto 2xTY agar  
14 plates supplemented with 1% glucose and 100 $\mu$ g/ml  
15 ampicillin. Colonies were allowed to grow overnight at  
16 30°C. A 100-fold concentrated PDCP stock was prepared  
17 from a 200ml amplified culture of these bacteria as  
18 described above, and 0.5ml used in as second round of  
19 selection with 500ng biotinylated substance P. For this  
20 round 100 $\mu$ g/ml streptavidin was included in the wash  
21 buffer.

### 22 ELISA identification of binding clones.

23 Binding clones were identified by ELISA of '96  
24 individual PDCP cultures prepared as described in  
25 Example 3 from colonies recovered after the second  
26 round of selection. A Dynatech Immulon 4 ELISA plate  
27 was coated with 200ng/well streptavidin in 100 $\mu$ l /well  
28 PBS for 1 hour at 37°C. The plate was washed  
29 3x200 $\mu$ l/well PBS and incubated with 10ng/well  
30 biotinylated substance P in 100 $\mu$ l /well PBS for 30  
31 minutes at 37°C The plate was washed 3x200 $\mu$ l/well PBS  
32 and blocked for 1 hour at 37°C with 200 $\mu$ l/well 2%  
33 Marvel non-fat milk powder/PBS and then washed  
34 2x200 $\mu$ l/well PBS. 50 $\mu$ l PDCP culture supernatant was  
35 added to each well containing 50 $\mu$ l/well 4% Marvel/PBS,  
36

1 and allowed to bind for 1 hour at ambient temperature.  
2 The plate was washed three times with 200 $\mu$ l/well  
3 PBS/0.1% Tween 20, then three times with 200 $\mu$ l/well  
4 PBS. Bound PDCPs were detected with 100 $\mu$ l/well, 1:5000  
5 diluted anti-M13-HRP conjugate (Pharmacia) in 2%  
6 Marvel/PBS for 1 hour at ambient temperature and the  
7 plate washed six times as above. The plate was  
8 developed for 10 minutes at ambient temperature with  
9 100 $\mu$ l/well freshly prepared TMB (3,3',5,5'-  
10 Tetramethylbenzidine) substrate buffer (0.005% H<sub>2</sub>O<sub>2</sub>,  
11 0.1mg/ml TMB in 24mM citric acid/52mM sodium phosphate  
12 buffer pH 5.2). The reaction was stopped with  
13 100 $\mu$ l/well 12.5% H<sub>2</sub>SO<sub>4</sub> and read at 450nm. Out of 96  
14 clones tested, 10 gave signals greater than twice  
15 background (background = 0.05).

#### 16 17 Characterization of a binding clone.

18 A 50-fold concentrated PDCP stock was prepared from a  
19 100ml amplified culture of a single ELISA positive  
20 clone as described above. 10 $\mu$ l per well of this stock  
21 was tested in ELISA as described above for binding to  
22 streptavidin, streptavidin-biotinylated-substance P and  
23 streptavidin-biotinylated-CGRP (N-terminal  
24 biotinylated). Binding was only observed in  
25 streptavidin-biotinylated-substance P coated wells  
26 indicating that binding was specific. In addition,  
27 binding to streptavidin-biotinylated substance P was  
28 completely inhibited by incubating the PDCP with 1 $\mu$ g/ml  
29 free substance P (see Figure 8). The scFv VH (SEQ ID  
30 Nos 15 and 16) and VL (SEQ ID Nos 17 and 18) DNA and  
31 amino acid sequence was determined by DNA sequencing  
32 with oligonucleotides M13REV (SEQ ID No27) and ORSEQFOR  
33 (SEQ ID No 36) and is shown in Figure 9.

34  
35 The results indicate that target binding activities can  
36 be isolated from PDCP display libraries of human scFv

1 fragments.

2

3 **Example 10**

4 In another example the invention provides methods for  
5 screening a DNA library whose members require more than  
6 one chain for activity, as required by, for example,  
7 antibody Fab fragments for ligand binding. To increase  
8 the affinity of an antibody of known heavy and light  
9 chain sequence, libraries of unknown light chains  
10 co-expressed with a known heavy chain are screened for  
11 higher affinity antibodies. The known heavy chain  
12 antibody DNA sequence is joined to a nucleotide  
13 sequence encoding a oestrogen receptor DNA binding  
14 domain in a phage vector which does not contain the  
15 oestrogen receptor HRE sequence. The antibody DNA  
16 sequence for the known heavy chain (VH and CH1) gene is  
17 inserted in the 5' region of the oestrogen receptor DBD  
18 DNA, behind an appropriate promoter and translation  
19 sequences and a sequence encoding a signal peptide  
20 leader directing transport of the downstream fusion  
21 protein to the periplasmic space. The library of  
22 unknown light chains (VL and CL) is expressed  
23 separately from a phagemid expression vector which also  
24 contains the oestrogen receptor HRE sequence. Thus when  
25 both heavy and light chains are expressed in the same  
26 host cell, following infection with the phage  
27 containing the heavy chain-DBD fusion, the light chain  
28 phagemid vector is preferentially packaged into mature  
29 phage particles as single stranded DNA, which is bound  
30 by the heavy chain-DBD fusion protein during the  
31 packaging process. The light chain proteins are  
32 transported to the periplasm where they assemble with  
33 the heavy chain that is fused to the DBD protein as it  
34 exits the cell on the PDCP. In this example the DBD  
35 fusion protein and the HRE DNA sequences are not  
36 encoded on the same vector, the unknown peptide

- 1 sequences are present on the same vector as the HRE
- 2 sequence. Peptide display carrier packages (PDCP) which
- 3 encode the protein of interest can then be selected by
- 4 means of a ligand specific for the antibody.



Table 1 (i) Oligonucleotide primers used for human scFv library construction

## cDNA synthesis primers

IgMCDNAFOR	TGGAAGAGGCACGTTCTTTTCTTT
IgDCDNAFOR	CTCCTTCTTACTCTTGCTGGCGGT
IgKCDNAFOR	AGACTCTCCCCTGTTGAAGCTCTT
IgλCDNAFOR	TGAAGATTCTGTAGGGGCCACTGTCTT

## JHFOR primers

JH1-2FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCAGGGTGCC
JH3FOR	TGAACCGCCTCCACCTGAAGAGACGGTGACCATTGTCCC
JH4-5FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCAGGGTTCC
JH6FOR	TGAACCGCCTCCACCTGAGGAGACGGTGACCGTGGTCCC

## VH family BAK primers

VH1BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTGCAGCTGGTGCAGTCTGG
VH2BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTCAACTTAAGGGAGTCTGG
VH3BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTGCAGCTGGTGGAGTCTGG
VH4BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTGCAGCTGCAGGAGTCCGG
VH5BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTGCAGCTGTTGCAGTCTGC
VH6BAK	TTTTTGCCCCAGCCGGCCATGGCCCAGGTACAGCTGCAGCAGTCAGG

## Light chain FOR primers

SCFVKFOR	TTATTCGCGGCCGCCTAAACAGAGGCAGTTCCAGATTTC
SCFVλFOR	GTCACTTGCGGCCGCCTACAGTGTGGCCTTGTTGGCTTG

## VK family BAK primers

VK1BAK	TCTGGCGGTGGCGGATCGGACATCCAGATGACCCAGTCTCC
VK2BAK	TCTGGCGGTGGCGGATCGGATGTTGTGATGACTCAGTCTCC
VK3BAK	TCTGGCGGTGGCGGATCGGAAATTGTGTTGACGCAGTCTCC
VK4BAK	TCTGGCGGTGGCGGATCGGACATCGTGATGACCCAGTCTCC
VK5BAK	TCTGGCGGTGGCGGATCGGAAACGACACTCACGCAGTCTCC
VK6BAK	TCTGGCGGTGGCGGATCGGAAATTGTGCTGACTCAGTCTCC

## JK FOR primers

JK1FOR	TTCTCGTGCGGCCGCTAACGTTTGATTTCACCTTGGTCCC
JK2FOR	TTCTCGTGCGGCCGCTAACGTTTGATCTCCAGCTTGGTCCC
JK3FOR	TTCTCGTGCGGCCGCTAACGTTTGATATCCACTTTGGTCCC
JK4FOR	TTCTCGTGCGGCCGCTAACGTTTGATCTCCACCTTGGTCCC
JK5FOR	TTCTCGTGCGGCCGCTAACGTTTAACTCTCCAGTCGTGTCCC

## VA family BAK primers

VA1BAK	TCTGGCGGTGGCGGATCGCAGTCTGTGTTGACGCAGCCGCC
VA2BAK	TCTGGCGGTGGCGGATCGCAGTCTGCCCTGACTCAGCCTGC

Table 1 (ii) Oligonucleotide primers used for human scFv library construction

VA3aBAK	TCTGGCGGTGGCGGATCGTCCTATGTGCTGACTCAGCCACC
VA3bBAK	TCTGGCGGTGGCGGATCGTCTTCTGAGCTGACTCAGGACCC
VA4BAK	TCTGGCGGTGGCGGATCGCACGTTATACTGACTCAACCGCC
VA5BAK	TCTGGCGGTGGCGGATCGCAGGCTGTGCTCACTCAGCCGTC
VA6BAK	TCTGGCGGTGGCGGATCGAATTTTATGCTGACTCAGCCCCA

## JA primers

JA1FOR	TTCTCGTGCGGCCGCTAACCTAGGACGGTGACCTTGGTCCC
JA2-3FOR	TTCTCGTGCGGCCGCTAACCTAGGACGGTCAGCTTGGTCCC
JA4-5FOR	TTCTCGTGCGGCCGCTAACCTAAAACGGTGAGCTGGGTCCC

## Linker primers

LINKAMP3	CGATCCGCCACCGCCAGA
LINKAMP5	GTCTCCTCAGGTGGAGGC
LINKAMP3T	CGATCCGCCACCGCCAGAGCCACCTCCGCCTGAACCGCCTCCACCTGAGGAGAC

## 1 Claims

- 2
- 3 1. A peptide display carrier package (PDCP), said
- 4 package comprising a recombinant polynucleotide-
- 5 chimeric protein complex wherein the chimeric
- 6 protein has a nucleotide binding portion and a
- 7 target peptide portion, wherein said recombinant
- 8 polynucleotide comprises a nucleotide sequence
- 9 motif which is specifically bound by said
- 10 nucleotide binding portion, and wherein at least
- 11 the chimeric protein-encoding portion of the
- 12 recombinant polynucleotide not bound by the
- 13 chimeric protein nucleotide binding portion is
- 14 protected by a binding moiety.
- 15
- 16 2. A peptide display carrier package (PDCP) as claimed
- 17 in Claim 1, wherein said chimeric protein-encoding
- 18 portion of the recombinant polynucleotide not bound
- 19 by the chimeric protein nucleotide binding portion
- 20 is protected by a non-sequence-specific protein.
- 21
- 22 3. A peptide display carrier package (PDCP) as claimed
- 23 in Claim 2, wherein said non-sequence-specific
- 24 protein is a viral coat protein.
- 25
- 26 4. A peptide display carrier package (PDCP) as claimed
- 27 in any one of Claims 1 to 3, wherein said target
- 28 peptide portion is displayed externally on the
- 29 package.
- 30
- 31 5. A peptide display carrier package (PDCP) as claimed
- 32 in any one of Claims 1 to 4 wherein said
- 33 recombinant polynucleotide includes a linker
- 34 sequence between the nucleotide sequence encoding
- 35 the nucleotide binding portion and the nucleotide
- 36 sequence encoding the target peptide portion.

- 1     6. A peptide display carrier package (PDCP) as claimed  
2       in any one of Claims 1 to 5 wherein said  
3       recombinant polynucleotide has two or more  
4       nucleotide sequence motifs each of which can be  
5       bound by the nucleotide binding portion of the  
6       chimeric protein.  
7
- 8     7. A peptide display carrier package (PDCP) as claimed  
9       in any one of Claims 1 to 6 wherein said nucleotide  
10      binding portion is a DNA binding domain of an  
11      oestrogen or progesterone receptor.  
12
- 13    8. A peptide display carrier package (PDCP) as claimed  
14      in any one of Claims 1 to 7 wherein said  
15      recombinant polynucleotide is bound to said  
16      chimeric protein as single stranded DNA.  
17
- 18    9. A peptide display carrier package (PDCP) as claimed  
19      in any one of Claims 1 to 8 wherein said target  
20      peptide portion is located at the N and/or C  
21      terminal of the chimeric protein.  
22
- 23    10. A peptide display carrier package (PDCP) as claimed  
24       in any one of Claims 1 to 9 which is produced in a  
25       host cell transformed with said recombinant  
26       polynucleotide and extruded therefrom without lysis  
27       of the host cell.  
28
- 29    11. A recombinant polynucleotide comprising a  
30       nucleotide sequence encoding a chimeric protein  
31       having a nucleotide binding portion operably linked  
32       to a target peptide portion, wherein said  
33       polynucleotide includes a specific nucleotide  
34       sequence motif which is bound by the nucleotide  
35       binding portion of said chimeric protein and  
36       further encoding a non-sequence-specific nucleotide

1 binding protein.

2

3 12. A recombinant polynucleotide as claimed in Claim 11  
4 wherein said non-sequence-specific nucleotide  
5 binding protein is a viral coat protein.

6

7 13. A recombinant polynucleotide as claimed in either  
8 one of Claims 11 and 12 which includes a linker  
9 sequence between the nucleotide sequence encoding  
10 the nucleotide binding portion and the nucleotide  
11 sequence encoding the target peptide portion.

12

13 14. A recombinant polynucleotide as claimed in any one  
14 of Claims 11 to 13 which has two or more nucleotide  
15 sequence motifs each of which can be bound by the  
16 nucleotide binding portion of the chimeric protein.

17

18 15. A recombinant polynucleotide as claimed in any one  
19 of Claims 11 to 14 wherein said nucleotide binding  
20 portion is a DNA binding domain of an oestrogen or  
21 progesterone receptor.

22

23 16. A recombinant polynucleotide as claimed in any one  
24 of Claims 11 to 15 wherein said recombinant  
25 polynucleotide is bound to said chimeric protein as  
26 single stranded DNA.

27

28 17. A genetic construct or set of genetic constructs  
29 which collectively comprises a polynucleotide  
30 having a sequence which includes:

- 31 i) a sequence encoding a nucleotide binding  
32 portion able to recognise and bind to a  
33 specific sequence motif;  
34 ii) the sequence motif recognised and bound by the  
35 nucleotide binding portion encoded by (i);  
36 iii) a restriction enzyme site which permits

1 insertion of a polynucleotide, said site being  
2 designed to operably link said polynucleotide  
3 to the sequence encoding the nucleotide  
4 binding portion so that expression of the  
5 operably linked polynucleotide sequences  
6 yields a chimeric protein; and  
7 iv) a sequence encoding a nucleotide binding  
8 protein which binds non-specifically to naked  
9 polynucleotide.

10  
11 18. A genetic construct or set of genetic constructs as  
12 claimed in Claim 17 wherein a linker sequence is  
13 located between the nucleotide sequence encoding  
14 the nucleotide binding portion and the site for  
15 insertion of the polynucleotide.

16  
17 19. A genetic construct or set of genetic constructs as  
18 claimed in either one of Claims 17 and 18 which  
19 includes a vector pDM12, pDM14 or pDM16, deposited  
20 at NCIMB under Nos 40970, 40971 and 40972  
21 respectively.

22  
23 20. A method of constructing a genetic library, said  
24 method comprising:

25  
26 a) constructing multiple copies of a recombinant  
27 vector comprising a polynucleotide sequence  
28 which encodes a nucleotide binding portion  
29 able to recognise and bind to a specific  
30 sequence motif;

31  
32 b) operably linking each said vector to a  
33 polynucleotide encoding a target polypeptide,  
34 such that expression of said operably linked  
35 vector results in expression of a chimeric  
36 protein comprising said target peptide and

1           said nucleotide binding portions; wherein said  
2           multiple copies of said operably linked  
3           vectors collectively express a library of  
4           target peptide portions;

5  
6           c)   transforming host cells with the vectors of  
7           step b);

8  
9           d)   culturing the host cells of step c) under  
10          conditions suitable for expression of said  
11          chimeric protein;

12  
13          e)   providing a recombinant polynucleotide  
14          comprising the nucleotide sequence motif  
15          specifically recognised by the nucleotide  
16          binding portion and exposing this  
17          polynucleotide to the chimeric protein of step  
18          d) to yield a polynucleotide-chimeric protein  
19          complex; and

20  
21          f)   causing production of a non-sequence-specific  
22          moiety able to bind to the non-protected  
23          portion of the polynucleotide encoding the  
24          chimeric protein to form a peptide display  
25          carrier package.

26

27       21. A method of screening a genetic library, said  
28       method comprising:

29

30          a)   exposing the polynucleotide members of said  
31          library to multiple copies of a genetic  
32          construct comprising a nucleotide sequence  
33          encoding a nucleotide binding portion able to  
34          recognise and bind to a specific sequence  
35          motif, under conditions suitable for the  
36          polynucleotides of said library each to be

- 1 individually ligated into one copy of said  
2 genetic construct, to create a library of  
3 recombinant polynucleotides;  
4  
5 b) exposing said recombinant polynucleotides to a  
6 population of host cells, under conditions  
7 suitable for transformation of said host cells  
8 by said recombinant polynucleotides;  
9  
10 c) selecting for transformed host cells;  
11  
12 d) exposing said transformed host cells to  
13 conditions suitable for expression of said  
14 recombinant polynucleotide to yield a chimeric  
15 protein; and  
16  
17 e) providing a recombinant polynucleotide  
18 comprising the nucleotide sequence motif  
19 specifically recognised by the nucleotide  
20 binding portion and exposing this  
21 polynucleotide to the chimeric protein of step  
22 d) to yield a polynucleotide-chimeric protein  
23 complex;  
24  
25 f) protecting any exposed portions of the  
26 polynucleotide in the complex of step e) to  
27 form a peptide display carrier package; and  
28  
29 g) screening said peptide display carrier package  
30 to select only those packages displaying a  
31 target peptide portion having the  
32 characteristics required.

- 33  
34 22. A method as claimed in Claim 21 wherein the peptide  
35 display package carrier is extruded from the host  
36 cell without lysis thereof.



- 1 23. A polynucleotide comprising a nucleotide sequence
- 2 substantially as set out in SEQ ID No. 15 or SEQ ID
- 3 No. 17.

## Figure 1

1/9

1

pel B

MET LYS TYR LEU LEU PRO TER ALA ALA ALA GLY LEU

AAGCTTGCATGCAAATTCCTATTTCAAGGAGACAGTCATAA ATG AAA TAC CTA TTG CCT ACG GCA GCC GGT GGA TTG  
 TTCGAACGTACGTTTAAGATAAAGTTCCTCTGTCACTATT TAC TTT ATG GAT AAC GGA TGC CGT CGG CGA CCT AAC

*Hin* dIII

77

*Sfi* I*Pst* I*Not* I

LEU LEU LEU ALA ALA GLN PRO ALA MET ALA GLU VAL GLN LEU GLN \*\*\* \*\* ALA ALA ALA  
 TTA TTA CTC GCG GCC CAG CCG GCC ATG GCC CAG GTG CAA CTG CAG TAA TAG GCG GCC GCA  
 AAT AAT CAG CCG CGG GTC GCG CGG TAC CGG GTC CAC GTC GAC GTC ATT ATC CCG CGG CGT

137

GLY GLY GLY GLY SER MET GLU SER ALA LYS GLU THR ARG TYR CYS ALA VAL CYS ASN ASP  
 GGG GGA GCA GGG TCC ATG GAA TCT GCG AAG GAG ACT CGC TAC TGT GCA GTG TGC AAT GAC  
 CCC OCT OCT CCC AGG TAC CTT AGA CCG TTC CTC TGA CCG ATG ACA CGT CAC ACG TTA CTG

197

TYR ALA SER GLY TYR HIS TYR GLY VAL TRP SER CYS GLU GLY CYS LYS ALA PHE PHE LYS  
 TAT GCT TCA GGC TAC CAT TAT GGA GTC TGG TCC TGT GAG GGC TGC AAG GCC TTC TTC AAG  
 ATA CGA AGT CCG ATG GTA ATA OCT CAG ACC AGG ACA CTC CCG ACG TTC CGG AAG AAG TTC

257

ARG SER ILE GLN GLY HIS ASN ASP TYR MET CYS PRO ALA THR ASN GLN CYS THR ILE ASP  
 ACA AGT ATT CAA GGA CAT AAC GAC TAT ATG TGT CCA GCC ACC AAC CAG TGC ACC ATT GAT  
 TCT TCA TAA GTT OCT GTA TTG CTG ATA TAC ACA GGT CCG TGG TTG GTC ACG TGG TAA CTA

317

Oestrogen receptor DBD

LYS ASN ARG ARG LYS SER CYS GLN ALA CYS ARG LEU ARG LYS CYS TYR GLU VAL GLY MET  
 AAA AAC AGG AGG AAG AGC TGC CAG GCC TGC CCG CTC CGT AAA TGC TAC GAA GTG GGA ATG  
 TTT TTG TCC TCC TTC TCG ACG GTC CCG ACG GCC GAG GCA TTT ACG ATG CTT CAC CCT TAC

377

MET LYS GLY GLY ILE ARG LYS ASP ARG ARG GLY GLY ARG MET LEU LYS HIS LYS ARG GLN  
 ATG AAA GGT GGG ATA CGA AAA GAC CGA AGA GGA GGG AGA ATG TTG AAA CAC AAG CCG CAG  
 TAC TTT CCA CCC TAT GCT TTT CTG GCT TCT CCT CCC TCT TAC AAC TTT GTG TTC GCG GTC

437

ARG ASP ASP GLY GLU GLY ARG GLY GLU VAL GLY SER \*\*\* \*\*

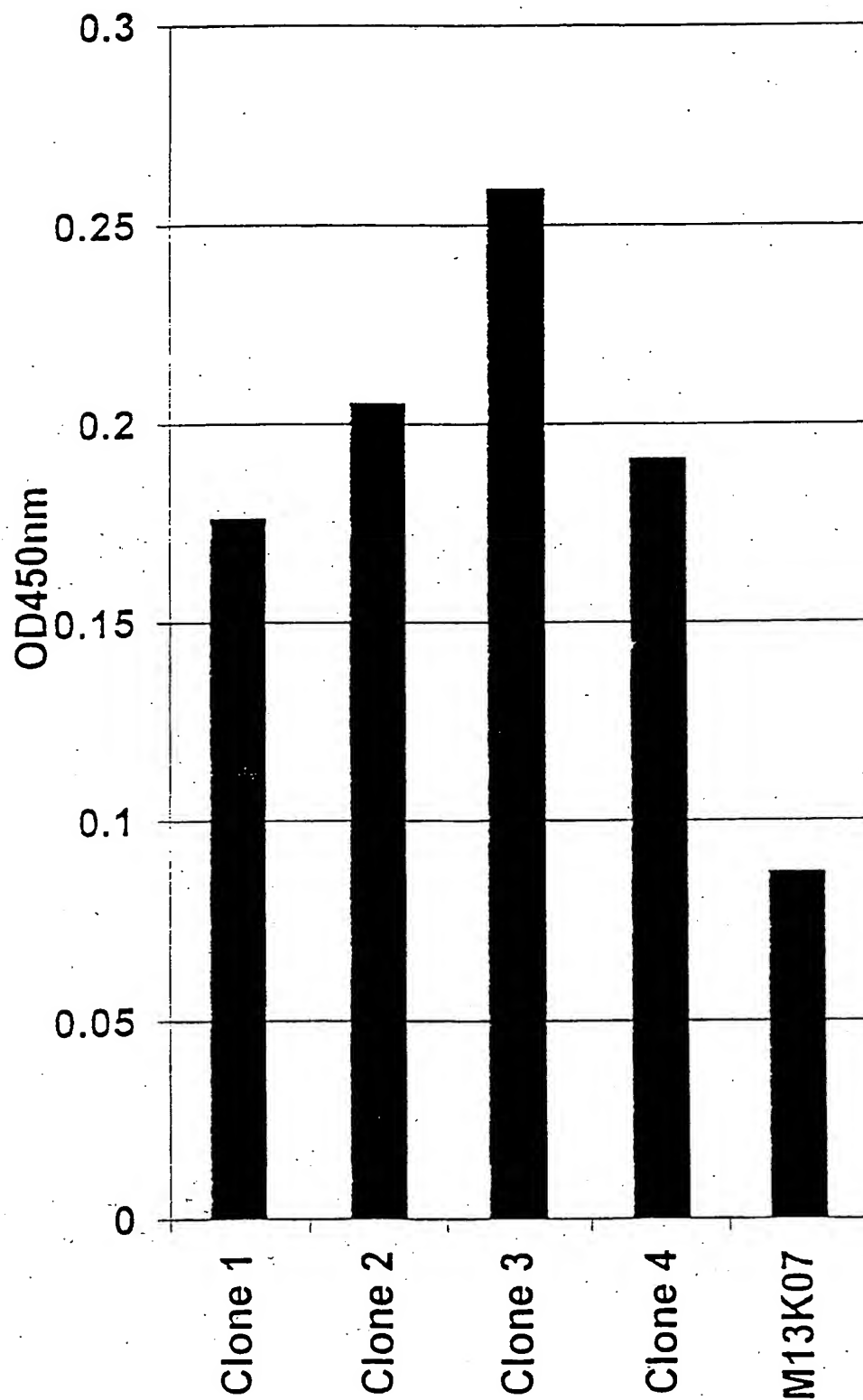
HRE

*Eco* RI

AGA GAT GAT GGG GAG GGC AGG GGT GAA GTG GGG TCT TGA TAA TCAGGTCAGAGTGACCTGAGCTAAAATAACACATTTCAG AATTC  
 TCT CTA CTA CCC CTC CCG TCC CCA CTT CAC CCC AGA ACT ATT AGTCCAGTCTCACTGGACTCGATTTTATTGTGTAAGTC TTAAG

2/9

Figure 2



## Figure 3

3/9

Human Igk constant region

K R T V A A P S V  
AAACGAACTGTGGCTGCACCATCTGTC

Clone #2

M A↓Q P T T R P G Q G T R L D I K R T V A A P S V  
ATGGCCCAGCCCACCACGCGTCCGGGCCAAGGGACACGACTGCACATTAAACGAACTGTGGCTGCACCATCTGTC

Clone #3

M A↓Q S H H A S G G G T K V E I K R T V A A P S V  
ATGGCCCAGTCCCACCACGCGTCCGGCGGAGGGACCAAGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTC

Human Igk constant region

F I F P P S D E Q L K S G T A S V V C L L N N F Y  
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

Clone #2

F I F P P S D E Q L K S G T A S V V C L L N N F Y  
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

Clone #3

F I F P P S D E Q L K S G T A S V V C L L N N F Y  
TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTAT

## Figure 4

4/9 pel B

1Hind III

MET LYS TYR LEU LEU PRO THR ALA ALA

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATAA ATG AAA TAC CTA TTG CCT ACG GCA GCC

TTCGAACGTA CGTTTAAGAT AAAGTTCCTC TGTCAGTATT TAC TTT ATG GAT AAC GGA TGC CGT CGG

68

Sfi I

Pst I

ALA GLY LEU LEU LEU LEU ALA ALA GLN PRO ALA MET ALA GLU VAL GLN LEU GLN \*\*\* \*\*

GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCC ATG GCC GAG GTG CAA CTG CAG TAA TAG

CGA CCT AAC AAT AAT GAG CGC CGG GTC GGC CGG TAC CGG CTC CAC GTT GAC GTC ATT ATC

128 Not I

ALA ALA ALA GLY GLY GLY GLY SER MET GLU SER ALA LYS GLU THR ARG TYR CYS ALA VAL

GCG GCC GCA GGG GGA GGA GGG TCC ATG GAA TCT GCC AAG GAG ACT CGC TAC TGT GCA GTG

CGC CGG CGT CCC CCT CCT CCC AGG TAC CTT AGA CGG TTC CTC TGA GCG ATG ACA CGT CAC

188

CYS ASN ASP TYR ALA SER GLY TYR HIS TYR GLY VAL TRP SER CYS GLU GLY CYS LYS ALA

TGC AAT GAC TAT GCT TCA GGC TAC CAT TAT GGA GTC TGG TCC TGT GAG GGC TGC AAG GCC

ACG TTA CTG ATA CGA AGT CCG ATG GTA ATA CCT CAG ACC AGG ACA CTC CCG ACG TTC CGG

248

PHE PHE LYS ARG SER ILE GLN GLY HIS ASN ASP TYR MET CYS PRO ALA THR ASN GLN CYS

TTC TTC AAG AGA AGT ATT CAA GGA CAT AAC GAC TAT ATG TGT CCA GCC ACC AAC CAG TGC

AAG AAG TTC TCT TCA TAA GTT CCT GTA TTG CTG ATA TAC ACA GGT CGG TGG TTG GTC ACG

308

THR ILE ASP LYS ASN ARG ARG LYS SER CYS GLN ALA CYS ARG LEU ARG LYS CYS TYR GLU

ACC ATT GAT AAA AAC AGG AGG AAG AGC TGC CAG GCC TGC CGG CTC CGT AAA TGC TAC GAA

TGG TAA CTA TTT TTG TCC TCC TTC TCG ACG GTC CGG ACG GCC GAG GCA TTT ACG ATG CTT

368

VAL GLY MET MET LYS GLY GLY ILE ARG LYS ASP ARG ARG GLY GLY ARG MET LEU LYS HIS

GTG GGA ATG ATG AAA GGT GGG ATA CGA AAA GAC CGA AGA GGA GGG AGA ATG TTG AAA CAC

CAC CCT TAC TAC TTT CCA CCC TAT GCT TTT CTG GCT TCT CCT CCC TCT TAC AAC TTT GTG

428

LYS ARG GLN ARG ASP ASP GLY GLU GLY ARG GLY GLU VAL GLY SER Ter Ter

HRE 1

AAG CGC CAG AGA GAT GAT GGG GAG GGC AGG GGT GAA GTG GGG TCT TGA TAA TCAGGTCAGAGT

TTC GCG GTC TCT CTA CTA CCC CTC CCG TCC CCA CTT CAC CCC AGA ACT ATT AGTCCAGTCTCA

491

HRE 1

Sal I

HRE 2

Eco RI

GACCTGAGCTAAAATAACACATTGAG GTCGAC TTGGGTCAGTCTGACCGGGACAAAGTTAATGTAACCTC GAATTC

CTGGACTCGATTTTATTGTGTAAGTC CAGCTG AACCCAGTCAGACTGGCCCTGTTTCAATTACATTGGAG CTTAAG

## Figure 5

5/9

pel B

1 *Hind*III

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATAA ATG AAA TAC CTA TTG CCT ACG GCA GCC  
 TTCGAACGTA CGTTTAAGAT AAAGTTCCTC TGTCAGTATT TAC TTT ATG GAT AAC GGA TGC CGT CGG  
 68

ALA GLY LEU LEU LEU LEU ALA ALA GLN PRO ALA MET ALA GLU MET GLU SER ALA LYS GLU  
 GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCA ATG GCC GAG ATG GAA TCT GCC AAG GAG  
 CGA CCT AAC AAT AAT GAG CGC CGG GTC GGC CGT TAC CGG CTC TAC CTT AGA CGG TTC CTC  
 128

THR ARG TYR CYS ALA VAL CYS ASN ASP TYR ALA SER GLY TYR HIS TYR GLY VAL TRP SER  
 ACT CGC TAC TGT GCA GTG TGC AAT GAC TAT GCT TCA GGC TAC CAT TAT GGA GTC TGG TCC  
 TGA GCG ATG ACA CGT CAC ACG TTA CTG ATA CGA AGT CCG ATG GTA ATA CCT CAG ACC AGG  
 188

CYS GLU GLY CYS LYS ALA PHE PHE LYS ARG SER ILE GLN GLY HIS ASN ASP TYR MET CYS  
 TGT GAG GGC TGC AAG GCC TTC TTC AAG AGA AGT ATT CAA GGA CAT AAC GAC TAT ATG TGT  
 ACA CTC CCG ACG TTC CGG AAG AAG TTC TCT TCA TAA GTT CCT GTA TTG CTG ATA TAC ACA  
 248

PRO ALA THR ASN GLN CYS THR ILE ASP LYS ASN ARG ARG LYS SER CYS GLN ALA CYS ARG  
 CCA GCC ACC AAC CAG TGC ACC ATT GAT AAA AAC AGG AGG AAG AGC TGC CAG GCC TGC CGG  
 GGT CGG TGG TTG GTC ACG TGG TAA CTA TTT TTG TCC TCC TTC TCG ACG GTC CGG ACG GCC  
 308

LEU ARG LYS CYS TYR GLU VAL GLY MET MET LYS GLY GLY ILE ARG LYS ASP ARG ARG GLY  
 CTC CGT AAA TGC TAC GAA GTG GGA ATG ATG AAA GGT GGG ATA CGA AAA GAC CGA AGA GGA  
 GAG GCA TTT ACG ATG CTT CAC CCT TAC TAC TTT CCA CCC TAT GCT TTT CTG GCT TCT CCT  
 368

GLY ARG MET LEU LYS HIS LYS ARG GLN ARG ASP ASP GLY GLU GLY ARG GLY GLU VAL GLY  
 GGG AGA ATG TTG AAA CAC AAG CGC CAG AGA GAT GAT GGG GAG GGC AGG GGT GAA GTG GGG  
 CCC TCT TAC AAC TTT GTG TTC GCG GTC TCT CTA CTA CCC CTC CCG TCC CCA CTT CAC CCC  
 428

	<i>Sfi</i> I	<i>Pst</i> I	<i>Not</i> I
SER	GLY GLY GLY GLY SER	ALA GLN PRO ALA LEU LEU GLN LEU	ALA ALA ALA TER
TCT	GGG GGA GGA GGG TCG	GCC CAG CCG GCC CTC CTG CAG CTG	GCG GCC GCA TAACTGATTG
AGA	CCC CCT CCT CCC AGC	CGG GTC GGC CGG GAG GAC GTC GAC	GCG CGG CGT ATTGACTAAC

489

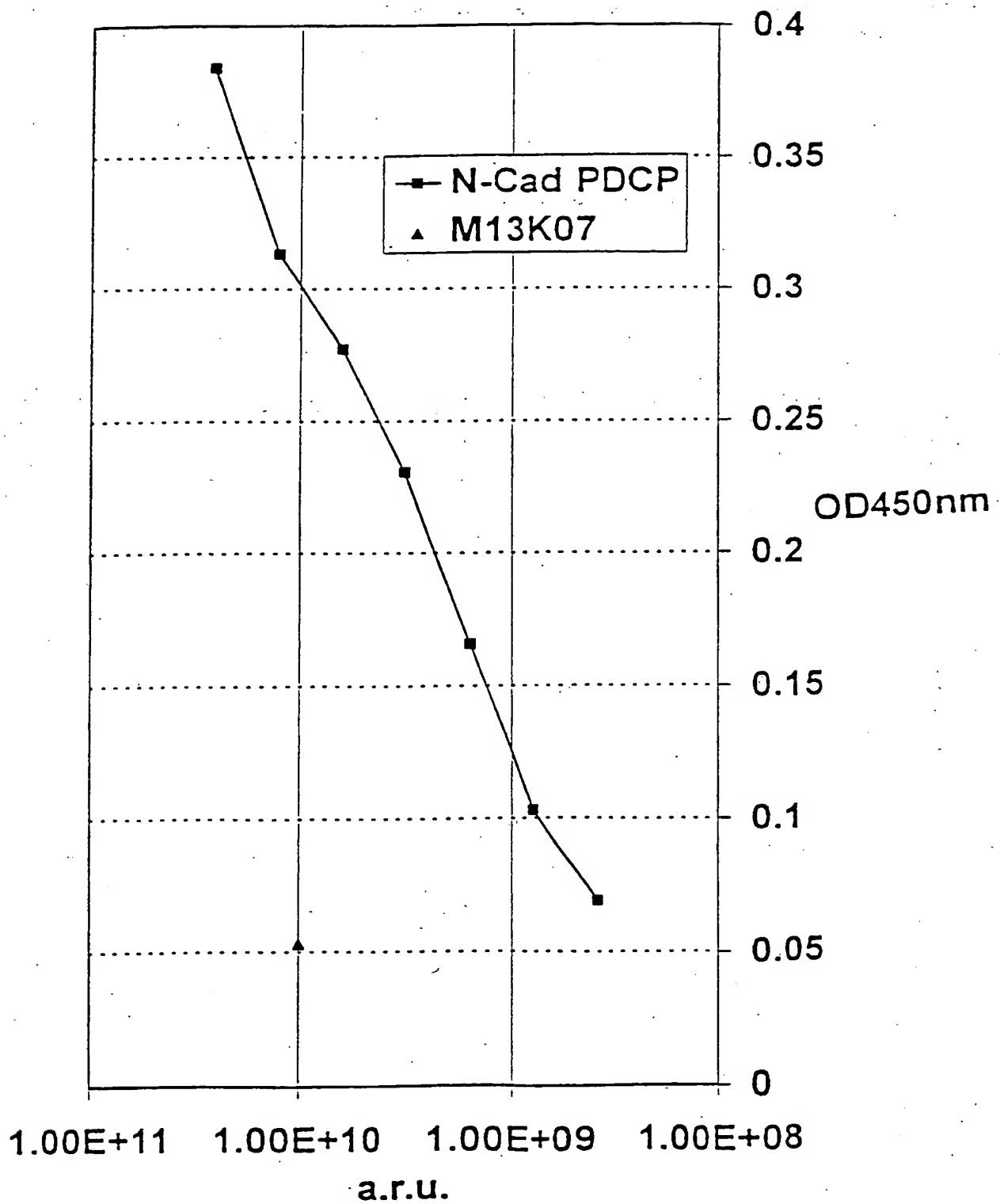
*Sal*I*Eco*RI

AGTCGACTTG GGTCAGTCTG ACCGGGACAA AGTTAATGTA ACCTC GAATTC  
 TCAGCTGAAC CCAGTCAGAC TGGCCCTGTT TCAATTACAT TGGAG CTTAAG

HRE

6/9

Figure 6



7/9

Figure 7

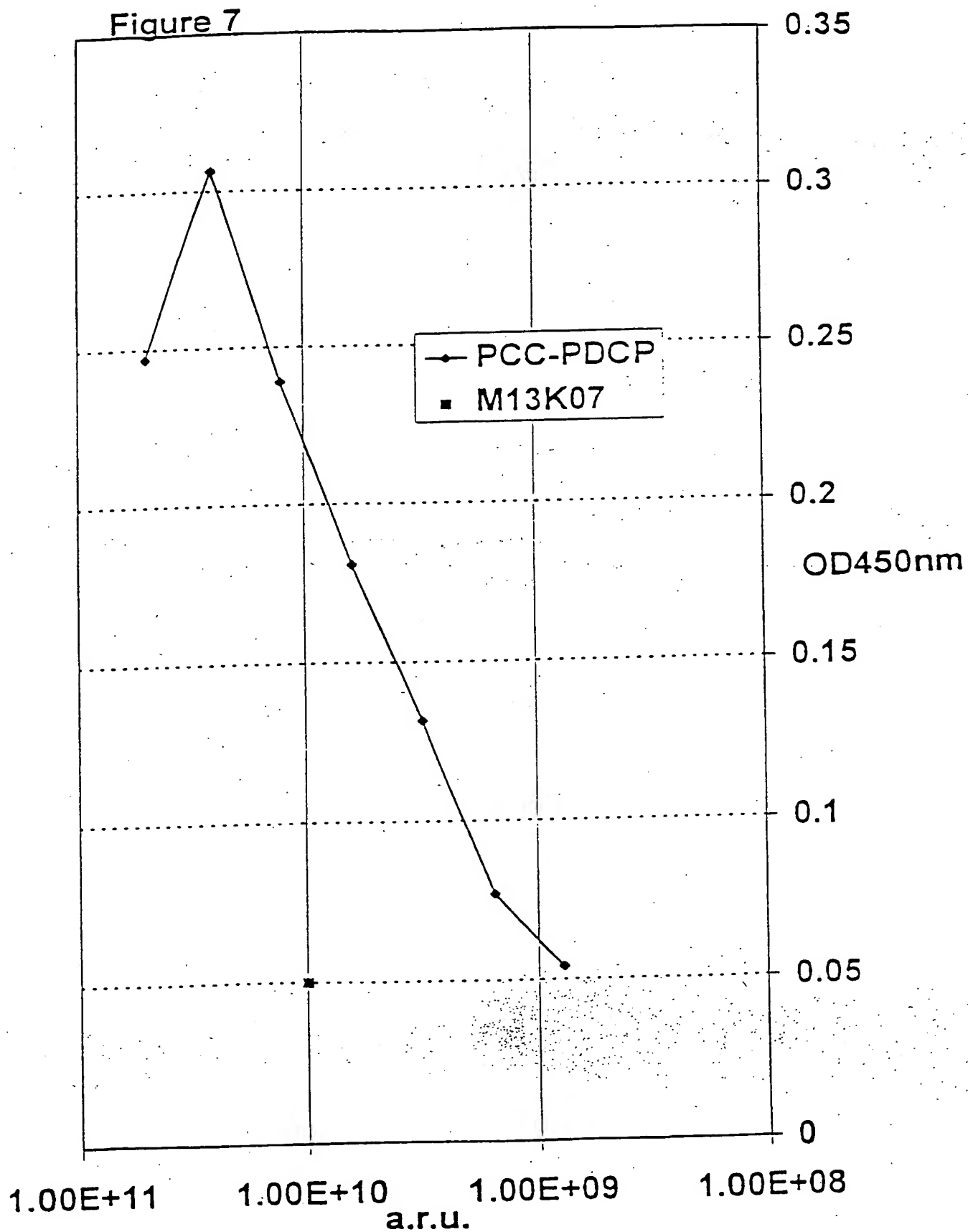
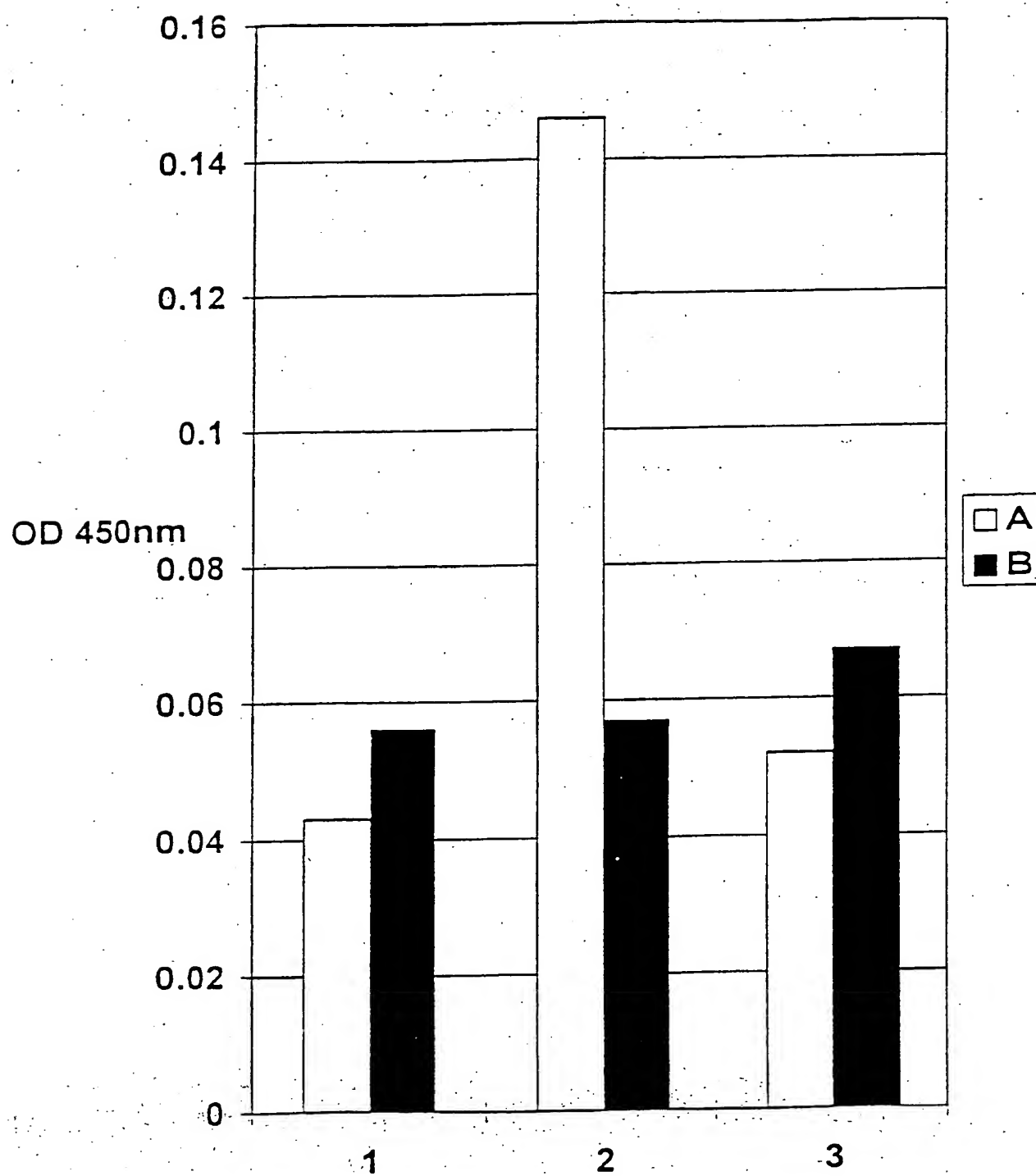




Figure 8

8/9



## HEAVY CHAIN

9/9

Figure 9

Q V Q L Q Q S G G G V V Q P G R S L  
 CAGGTACAGCTGCAGCAGTCAGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTG  
 GTCCATGTCGACGTCGTCAGTCCCCCTCCGCACCAGGTCCGACCCTCCAGGGAC  
 R L S C A A S G F P F S T Y G M H W  
 AGACTCTCCTGTGCAGCCTCGGGATTCCCCTTTAGTACTTATGGCATGCACTGG  
 TCTGAGAGGACACGTCGGAGCCCTAAGGGGAAATCATGAATACCGTACGTGACC  
 R Q A V P G K G L E W V A V I S Y D  
 CGCCAGGCTGTCCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGAT  
 GCGGTCCGACAGGGTCCGTTCCCCGACCTCACCCACCGTCAATATAGTATACTA  
 G S N K Y Y A D S V K G R F T I S R  
 GGAAGTAATAAATACTACGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGA  
 CCTTCATTATTTATGATGCGTCTGAGGCACTTCCCGGCTAAGTGGTAGAGGTCT  
 D N S K N T L Y L Q M N S L R A E D  
 GACAATTCCAAGAACACGTTGTATCTGCAAATGAACAGCCTGAGAGCTGAGGAC  
 CTGTTAAGGTTCTTGTGCAACATAGACGTTTACTTGTGCGACTCTCGACTCCTG  
 T A V Y Y C A R D L D P T R Y S S G  
 ACGGCTGTGTATTACTGTGCGAGAGATTTAGACCCCAACAGGTATAGCAGTGGC  
 TGCCGACACATAATGACACGCTCTCTAAATCTGGGGTGGTCCATATCGTCACCG  
 W D T D Y W G Q G H L V T V S S  
 TGGGACACTGACTACTGGGGCCAGGGGCACCTGGTCACTGTCTCCTCA  
 ACCCTGTGACTGATGACCCCGGTCCCCGTGGACCAGTGACAGAGGAGT

## LIGHT CHAIN

E T T L T Q S P G T L S L S P G E R  
 GAAACGACACTCACGCAGTCTCCAGGCACCCTGTCTTTGTCTCCGGGGGAAAGA  
 CTTTGTCTGTGAGTGCGTCAGAGGTCCGTGGGACAGAAACAGAGGCCCTTTCT  
 A T L S C R A S Q N I G S S S L A W  
 GCCACCCTCTCCTGCAGGGCCAGTCAGAATATTGGCAGCAGCTCCTTAGCCTGG  
 CGGTGGGAGAGGACGTCCCGGTCACTCTTATAACCGTCGTGAGGAATCGGACC  
 Y Q Q K P G Q A P R L L I Y G A S T  
 TACCAACAGAAACCTGGCCAGGCTCCCAGGCTCCTCATCTATGGTGCATCCACC  
 ATGGTTGTCTTTGGACCGGTCCGAGGGTCCGAGGAGTAGATAACCACGTAGGTGG  
 R A T G F S G S G S G T Q F T L T I  
 AGGGCCACTGGTTTTCAGTGGCAGTGGGTGAGGACACAATTCAGTCTCACCATC  
 TCCCGGTGACCAAAGTCACCGTCACCCAGTCCCTGTGTTAAGTGAGAGTGGTAG  
 I P A R S S L Q S E D F A V Y Y C Q  
 ATCCAGCCAGGAGCAGCCTGCAGTCTGAAGATTTTGCAGTTTATTACTGTGAG  
 TAGGGTCCGGTCTCGTCGGACGTCAGACTTCTAAACGTCAAATAATGACAGTC  
 Q Y N F W P F T F G P G T K L E I K  
 CAGTATAATTTCTGGCCATTCACTTTTGGCCCTGGGACCAAGCTGGAGATCAAA  
 GTCATATTAAAGACCGGTAAGTGAAACCGGGACCCTGGTTCGACCTCTAGTTT

R  
 CGT  
 GCA

1

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Rowett Research Services Limited
- (B) STREET: Greenburn
- (C) CITY: Buckburn ABERDEEN
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): AB21 9SB

(ii) TITLE OF INVENTION: Chimeric binding peptide library screening method

(iii) NUMBER OF SEQUENCES: 76

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) -DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCAGGTCAGA GTGACCTGAG CTAATAAC ACATTGAG

33

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AGTCCAGTCT CACTGGACTC GATTTTATTG TGTAAGTC

38

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 521 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 41..475

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATAA	ATG AAA TAC CTA TTG	55
	Met Lys Tyr Leu Leu	
	1 5	
CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCC ATG		103
Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met		
	10 15 20	
GCC CAA GTG CAG CTG CAG TAA TAG GCG GCC GCA GGG GGA GGA GGG TCC		151
Ala Gln Val Gln Leu Gln * * Ala Ala Ala Gly Gly Gly Gly Ser		
	25 30 35	
ATG GAA TCT GCC AAG GAG ACT CGC TAC TGT GCA GTG TGC AAT GAC TAT		199
Met Glu Ser Ala Lys Glu Thr Arg Tyr Cys Ala Val Cys Asn Asp Tyr		
	40 45 50	
GCT TCA GGC TAC CAT TAT GGA GTC TGG TCC TGT GAG GGC TGC AAG GCC		247
Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala		
	55 60 65	
TTC TTC AAG AGA AGT ATT CAA GGA CAT AAC GAC TAT ATG TGT CCA GCC		295
Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met Cys Pro Ala		
	70 75 80 85	
ACC AAC CAG TGC ACC ATT GAT AAA AAC AGG AGG AAG AGC TGC CAG GCC		343
Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala		
	90 95 100	
TGC CGG CTC CGT AAA TGC TAC GAA GTG GGA ATG ATG AAA GGT GGG ATA		391
Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys Gly Gly Ile		
	105 110 115	
CGA AAA GAC CGA AGA GGA GGG AGA ATG TTG AAA CAC AAG CGC CAG AGA		439
Arg Lys Asp Arg Arg Gly Gly Arg Met Leu Lys His Lys Arg Gln Arg		

120 125 130 135 140 145 150  
 GAT GAT GGG GAG GGC AGG GGT GAA GTG GGG TCT TGA TAATCAGGTC  
 Asp Asp Gly Glu Gly Arg Gly Glu Val Gly Ser  
 135 140 145 150  
 AGAGTGACCT GAGCTAAAAAT AACACATTCA GAATTC

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 145 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein  
(iii) FUNCTION DESCRIPTION:

(ii) MOLECULE TYPE: protein  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

(xi) SEQUENCE DETERMINATION

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala  
1 5 10 15  
Ala Gln Pro Ala Met Ala Gln Val Gln Leu Gln Ala Ala Ala  
20 25 30  
Gly Gly Gly Gly Ser Met Glu Ser Ala Lys Glu Thr Arg Tyr Cys Ala  
35 40 45  
Val Cys Asn Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys  
50 55 60  
Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp  
65 70 75 80  
Tyr Met Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg  
85 90 95  
Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met  
100 105 110  
Met Lys Gly Gly Ile Arg Lys Asp Arg Arg Gly Gly Arg Met Leu Lys  
115 120 125  
His Lys Arg Gln Arg Asp Asp Gly Glu Gly Arg Gly Glu Val Gly Ser  
130 135 140

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 102 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

4

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..102

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

AAA CGA ACT GTG GCT GCA CCA TCT GTC TTC ATC TTC CCG CCA TCT GAT      48
Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
                150                155                160

GAG CAG TTG AAA TCT GGA ACT GCC TCT GTT GTG TGC CTG CTG AAT AAC      96
Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
                165                170                175

TTC TAT      102
Phe Tyr

```

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
 1           5           10           15
Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
 20           25           30
Phe Tyr

```

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 150 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..150

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

																5	
ATG	GCC	CAG	CCC	ACC	ACG	CGT	CCG	GGC	CAA	GGG	ACA	CGA	CTG	GAC	ATT	48	
Met	Ala	Gln	Pro	Thr	Thr	Arg	Pro	Gly	Gln	Gly	Thr	Arg	Leu	Asp	Ile		
				40				45				50					
35																	
AAA	CGA	ACT	GTG	GCT	GCA	CCA	TCT	GTC	TTC	ATC	TTC	CCG	CCA	TCT	GAT	96	
Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp		
				55				60				65					
GAG	CAG	TTG	AAA	TCT	GGA	ACT	GCC	TCT	GTT	GTG	TGC	CTG	CTG	AAT	AAC	144	
Glu	Gln	Leu	Lys	Ser	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	Leu	Asn	Asn		
				70				75				80					
TTC	TAT															150	
Phe	Tyr																

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Ala Gln Pro Thr Thr Arg Pro Gly Gln Gly Thr Arg Leu Asp Ile  
1 5 10 15

Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp  
20 25 30

Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn  
35 40 45

Phe Tyr  
50

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 150 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..150

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

6

ATG	GCC	CAG	TCC	CAC	CAC	GCG	TCC	GGC	GGA	GGG	ACC	AAG	GTG	GAG	ATC	48
Met	Ala	Gln	Ser	His	His	Ala	Ser	Gly	Gly	Gly	Thr	Lys	Val	Glu	Ile	
				55					60					65		
AAA	CGA	ACT	GTG	GCT	GCA	CCA	TCT	GTC	TTC	ATC	TTC	CCG	CCA	TCT	GAT	96
Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp	
			70					75					80			
GAG	CAG	TTG	AAA	TCT	GGA	ACT	GCC	TCT	GTT	GTG	TGC	CTG	CTG	AAT	AAC	144
Glu	Gln	Leu	Lys	Ser	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	Leu	Asn	Asn	
		85					90				95					
TTC	TAT															150
Phe	Tyr															
	100															

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Met	Ala	Gln	Ser	His	His	Ala	Ser	Gly	Gly	Gly	Thr	Lys	Val	Glu	Ile	
1				5				10						15		
Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp	
			20					25					30			
Glu	Gln	Leu	Lys	Ser	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	Leu	Asn	Asn	
		35					40					45				
Phe	Tyr															
	50															

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 566 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 41..475

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AAGCTTGCAT	GCAAATTCTA	TTTCAAGGAG	ACAGTCATAA	ATG	AAA	TAC	CTA	TTG	55
------------	------------	------------	------------	-----	-----	-----	-----	-----	----



7

Met Lys Tyr Leu Leu  
55

CCT	ACG	GCA	GCC	GCT	GGA	TTG	TTA	TTA	CTC	GCG	GCC	CAG	CCG	GCC	ATG	103
Pro	Thr	Ala	Ala	Ala	Gly	Leu	Leu	Leu	Leu	Ala	Ala	Gln	Pro	Ala	Met	
				60					65						70	
GCC	GAG	GTG	CAA	CTG	CAG	TAA	TAG	GCG	GCC	GCA	GGG	GGA	GGA	GGG	TCC	151
Ala	Glu	Val	Gln	Leu	Gln	*	*	Ala	Ala	Ala	Gly	Gly	Gly	Gly	Ser	
			75					80						95		
ATG	GAA	TCT	GCC	AAG	GAG	ACT	CGC	TAC	TGT	GCA	GTG	TGC	AAT	GAC	TAT	199
Met	Glu	Ser	Ala	Lys	Glu	Thr	Arg	Tyr	Cys	Ala	Val	Cys	Asn	Asp	Tyr	
		90					95					100				
GCT	TCA	GGC	TAC	CAT	TAT	GGA	GTC	TGG	TCC	TGT	GAG	GCC	TGC	AAG	GCC	247
Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp	Ser	Cys	Glu	Gly	Cys	Lys	Ala	
	105					110					115					
TTC	TTC	AAG	AGA	AGT	ATT	CAA	GGA	CAT	AAC	GAC	TAT	ATG	TGT	CCA	GCC	295
Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His	Asn	Asp	Tyr	Met	Cys	Pro	Ala	
120					125					130					135	
ACC	AAC	CAG	TGC	ACC	ATT	GAT	AAA	AAC	AGG	AGG	AAG	AGC	TGC	CAG	GCC	343
Thr	Asn	Gln	Cys	Thr	Ile	Asp	Lys	Asn	Arg	Arg	Lys	Ser	Cys	Gln	Ala	
				140					145					150		
TGC	CGG	CTC	CGT	AAA	TGC	TAC	GAA	GTG	GGA	ATG	ATG	AAA	GGT	GGG	ATA	391
Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val	Gly	Met	Met	Lys	Gly	Gly	Ile	
			155				160						165			
CGA	AAA	GAC	CGA	AGA	GGA	GGG	AGA	ATG	TTG	AAA	CAC	AAG	CGC	CAG	AGA	439
Arg	Lys	Asp	Arg	Arg	Gly	Gly	Arg	Met	Leu	Lys	His	Lys	Arg	Gln	Arg	
		170					175					180				
GAT	GAT	GGG	GAG	GGC	AGG	GGT	GAA	GTG	GGG	TCT	TGA	TAATCAGGTC				485
Asp	Asp	Gly	Glu	Gly	Arg	Gly	Glu	Val	Gly	Ser	*					
	185					190				195						
AGAGTGACCT	GAGCTAAAAT	AACACATTCA	GGTCGACTTG	GGTCAGTCTG	ACCGGGACAA											545
AGTTAATGTA	ACCTCGAATT	C														566

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 145 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met	Lys	Tyr	Leu	Leu	Pro	Thr	Ala	Ala	Ala	Gly	Leu	Leu	Leu	Leu	Ala	
1				5					10						15	
Ala	Gln	Pro	Ala	Met	Ala	Glu	Val	Gln	Leu	Gln	*	*	*	Ala	Ala	Ala
			20					25						30		
Gly	Gly	Gly	Gly	Ser	Met	Glu	Ser	Ala	Lys	Glu	Thr	Arg	Tyr	Cys	Ala	
		35					40					45				

8

Val Cys Asn Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys  
 50 55 60  
 Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp  
 65 70 75 80  
 Tyr Met Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg  
 85 90 95  
 Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met  
 100 105 110  
 Met Lys Gly Gly Ile Arg Lys Asp Arg Arg Gly Gly Arg Met Leu Lys  
 115 120 125  
 His Lys Arg Gln Arg Asp Asp Gly Glu Gly Arg Gly Glu Val Gly Ser  
 130 135 140

145

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 539 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 41..481

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AAGCTTGCAT GCAAATTCTA TTTCAAGGAG ACAGTCATAA ATG AAA TAC CTA TTG 55  
 Met Lys Tyr Leu Leu 150  
 CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCG GCC CAG CCG GCA ATG 103  
 Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met 165  
 GCC GAG ATG GAA TCT GCC AAG GAG ACT CGC TAC TGT GCA GTG TGC AAT 151  
 Ala Glu Met Glu Ser Ala Lys Glu Thr Arg Tyr Cys Ala Val Cys Asn 180  
 GAC TAT GCT TCA GGC TAC CAT TAT GGA GTC TGG TCC TGT GAG GGC TGC 199  
 Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys 195  
 AAG GCC TTC TTC AAG AGA AGT ATT CAA GGA CAT AAC GAC TAT ATG TGT 247  
 Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Met Cys

200	205	210	
CCA GCC ACC AAC CAG TGC ACC ATT GAT AAA AAC AGG AGG AAG AGC TGC			295
Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys			
215	220	225	230
CAG GCC TGC CGG CTC CGT AAA TGC TAC GAA GTG GGA ATG ATG AAA GGT			343
Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Met Lys Gly			
	235	240	245
GGG ATA CGA AAA GAC CGA AGA GGA GGG AGA ATG TTG AAA CAC AAG CGC			391
Gly Ile Arg Lys Asp Arg Arg Gly Gly Arg Met Leu Lys His Lys Arg			
	250	255	260
CAG AGA GAT GAT GGG GAG GGC AGG GGT GAA GTG GGG TCT GGG GGA GGA			439
Gln Arg Asp Asp Gly Glu Gly Arg Gly Glu Val Gly Ser Gly Gly Gly			
	265	270	275
GGG TCG GCC CAG CCG GCC CTC CTG CAG CTG GCG GCC GCA TAA			481
Gly Ser Ala Gln Pro Ala Leu Leu Gln Leu Ala Ala			
280	285	290	
CTGATTGAGT CGACTTGGGT CAGTCTGACC GGGACAAAGT TAATGTAACC TCGAATTC			539

## (2) INFORMATION FOR SEQ ID NO: 14:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 147 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met	Lys	Tyr	Leu	Leu	Pro	Thr	Ala	Ala	Ala	Gly	Leu	Leu	Leu	Leu	Ala
1				5					10					15	
Ala	Gln	Pro	Ala	Met	Ala	Glu	Met	Glu	Ser	Ala	Lys	Glu	Thr	Arg	Tyr
		20						25					30		
Cys	Ala	Val	Cys	Asn	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp
		35					40					45			
Ser	Cys	Glu	Gly	Cys	Lys	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His
	50					55					60				
Asn	Asp	Tyr	Met	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp	Lys	Asn
65				70						75				80	
Arg	Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val
			85						90					95	
Gly	Met	Met	Lys	Gly	Gly	Ile	Arg	Lys	Asp	Arg	Arg	Gly	Gly	Arg	Met
			100					105					110		
Leu	Lys	His	Lys	Arg	Gln	Arg	Asp	Asp	Gly	Glu	Gly	Arg	Gly	Glu	Val
		115					120				125				
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Gln	Pro	Ala	Leu	Leu	Gln	Leu	Ala
130						135					140				

10

Ala Ala  
145

## (2) INFORMATION FOR SEQ ID NO: 15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..372

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CAG GTA CAG CTG CAG CAG TCA GGG GGA GGC GTG GTC CAG CCT GGG AGG	43
Gln Val Gln Leu Gln Gln Ser Gly Gly Gly Val Val Gln Pro Gly Arg	
150 155 160	
TCC CTG AGA CTC TCC TGT GCA GCC TCG GGA TTC CCC TTT AGT ACT TAT	96
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Ser Thr Tyr	
165 170 175	
GGC ATG CAC TGG CGC CAG GCT GTC CCA GGC AAG GGG CTG GAG TGG GTG	144
Gly Met His Trp Arg Gln Ala Val Pro Gly Lys Gly Leu Glu Trp Val	
180 185 190	
GCA GTT ATA TCA TAT GAT GGA AGT AAT AAA TAC TAC GCA GAC TCC GTG	192
Ala Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val	
200 205 210	
AAG GGC CGA TTC ACC ATC TCC AGA GAC AAT TCC AAG AAC ACG TTG TAT	240
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr	
215 220 225	
CTG CAA ATG AAC AGC CTG AGA GCT GAG GAC ACG GCT GTG TAT TAC TGT	298
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys	
230 235 240	
GCG AGA GAT TTA GAC CCC ACC AGG TAT AGC AGT GGC TGG GAC ACT GAC	336
Ala Arg Asp Leu Asp Pro Thr Arg Tyr Ser Ser Gly Trp Asp Thr Asp	
245 250 255	
TAC TGG GGC CAG GGG CAC CTG GTC ACT GTC TCC TCA	372
Tyr Trp Gly Gln Gly His Leu Val Thr Val Ser Ser	
260 265 270	

## (2) INFORMATION FOR SEQ ID NO: 16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids

11

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```

Gln Val Gln Leu Gln Gln Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1           5           10
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Ser Thr Tyr
          20           25           30
Gly Met His Trp Arg Gln Ala Val Pro Gly Lys Gly Leu Glu Trp Val
          35           40           45
Ala Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val
          50           55           60
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
          65           70           75           80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
          85           90           95
Ala Arg Asp Leu Asp Pro Thr Arg Tyr Ser Ser Gly Trp Asp Thr Asp
          100          105          110
Tyr Trp Gly Gln Gly His Leu Val Thr Val Ser Ser
          115          120

```

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 327 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..327

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```

GAA ACG ACA CTC ACG CAG TCT CCA GGC ACC CTG TCT TTG TCT CCG GGG      48
Glu Thr Thr Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
125          130          135          140
GAA AGA GCC ACC CTC TCC TGC AGG GCC AGT CAG AAT ATT GGC AGC AGC      96
Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Asn Ile Gly Ser Ser
          145          150          155
TCC TTA GCC TGG TAC CAA CAG AAA CCT GGC CAG GCT CCC AGG CTC CTC      144
Ser Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu

```

																12	
160								165				170					
ATC	TAT	GGT	GCA	TCC	ACC	AGG	GCC	ACT	GGT	ATC	CCA	GCC	AGG	TTC	AGT	192	
Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala	Thr	Gly	Ile	Pro	Ala	Arg	Phe	Ser		
							180								185		
																240	
GGC	AGT	GGG	TCA	GGG	ACA	CAA	TTC	ACT	CTC	ACC	ATC	AGC	AGC	CTG	CAG		
Gly	Ser	Gly	Ser	Gly	Thr	Gln	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu	Gln		
							190								200		
																288	
TCT	GAA	GAT	TTT	GCA	GTT	TAT	TAC	TGT	CAG	CAG	TAT	AAT	TTC	TGG	CCA		
Ser	Glu	Asp	Phe	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	Tyr	Asn	Phe	Trp	Pro		
205								210								215	220
																327	
TTC	ACT	TTT	GGC	CCT	GGG	ACC	AAG	CTG	GAG	ATC	AAA	CGT					
Phe	Thr	Phe	Gly	Pro	Gly	Thr	Lys	Leu	Glu	Ile	Lys	Arg					
				225								230					

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Glu	Thr	Thr	Leu	Thr	Gln	Ser	Pro	Gly	Thr	Leu	Ser	Leu	Ser	Pro	Gly
1				5				10						15	
Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg	Ala	Ser	Gln	Asn	Ile	Gly	Ser	Ser
			20					25					30		
Ser	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Ala	Pro	Arg	Leu	Leu
		35					40					45			
Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala	Thr	Gly	Ile	Pro	Ala	Arg	Phe	Ser
	50					55					60				
Gly	Ser	Gly	Ser	Gly	Thr	Gln	Phe	Thr	Leu	Thr	Ile	Ser	Ser	Leu	Gln
65					70					75					80
Ser	Glu	Asp	Phe	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	Tyr	Asn	Phe	Trp	Pro
				85					90					95	
Phe	Thr	Phe	Gly	Pro	Gly	Thr	Lys	Leu	Glu	Ile	Lys	Arg			
			100					105							

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

13

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TTTTCTGCAG TAATAGGCGG CCGCAGGGGG AGGAGGGTCC ATCGAAGGTC GCGAAGCAGA 60  
GACTGTTGAA AG 72

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TTTTGAATTC TTATTAACCA CCGAACTGCG GGTGACGCCA AGCGCTTGCG GCCGTTAAGA 60  
CTCCTTATTA CGCAG 75

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

AAAAGCGGCC GCACTGGCCT GAGAGANNNN NN 32

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs

14

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TCGACCCACG CGTCCG

16

- (2) INFORMATION FOR SEQ ID-NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GGGTGCCGAG GC

12

- (2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 72 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

AAAAGAATTC TGAATGTGTT ATTTAGCTC AGGTCACCTCT GACCTGATTA TCAAGACCCC

60



ACTTCACCCC CT

72

## (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 46 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AAAAGCGGCC GCAGGGGGAG GAGGGTCCAT GGAATCTGCC AAGGAG

46

## (2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GTAAAACGAC GGCCAGT

17

## (2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GGATAACAAT TTCACACAGG

20

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

AAAGCGGCCG CACTGGCCTG AGAGA

25

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

AAAAGGCCCA GCCGGCCATG GCCCAGCCCA CCACGCGTCC G

41

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

17

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

AAAAGGCCCA GCCGGCCATG GCCCAGTCCC ACCACGCGTC CG

42

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

AAAAGGCCCA GCCGGCCATG GCCCAGTACC CACCACGCGT CCG

43

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

AAAAGAATTC GAGGTTACAT TAACTTTGTT CCGGTCAGAC TGACCCAAGT CGACCTGAAT

60

GTGTTATTTT AG

72

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

18

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CCTTGGCAGA TTCCATCTCG GCCATTGCCG GC

32

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CCGGCAATGG CCGAGATGGA ATCTGCCAAG G

31

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 86 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

TTTTGTGAC TCAATCAGTT ATGCGGCCGC CAGCTGCAGG AGGGCCGGCT GGGCCGACCC

60

TCCTCCCCCA GACCCCACTT CACCCC

86

19

## (2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

TGTTGAAACA CAAGCGCCAG

20

## (2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

TGGAAGAGGC ACGTTCTTTT CTTT

24

## (2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

20

CTCCTTCTTA CTCTTGCTGG CGGT

24

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

AGACTCTCCC CTGTTGAAGC TCTT

24

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

TGAAGATTCT GTAGGGGCCA CTGTCTT

27

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

TGAACCGCCT CCACCTGAGG AGACGGTGAC CAGGGTGCC

39

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TGAACCGCCT CCACCTGAAG AGACGGTGAC CATTGTCCC

39

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

TGAACCGCCT CCACCTGAGG AGACGGTGAC CAGGGTTCC

39

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

22

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

TGAACCGCCT CCACCTGAGG AGACGGTGAC CGTGGTCCC

39

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

TTTTTGCCCC AGCCGGCCAT GGCCAGGTG CAGCTGGTGC AGTCTGG

47

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

TTTTTGCCCC AGCCGGCCAT GGCCAGGTC AACTTAAGGG AGTCTGG

47

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"



23

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

TTTTTGGCCC AGCCGGCCAT GGCCGAGGTG CAGCTGGTGG AGTCTGG

47

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

TTTTTGGCCC AGCCGGCCAT GGCCGAGGTG CAGCTGCAGG AGTCGGG

47

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

TTTTTGGCCC AGCCGGCCAT GGCCGAGGTG CAGCTGTTGC AGTCTGC

47

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

TTTTTGGCCC AGCCGGCCAT GCCCCAGGTA CAGCTGCAGC AGTCAGG

47

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

TTATTCGCGG CCGCCTAAAC AGAGGCAGTT CCAGATTTC

39

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

GTCACCTGCG GCCGCCTACA GTGTGGCCTT GTTGGCTTG

39

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

25

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

TCTGGCGGTG GCGGATCGGA CATCCAGATG ACCCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

TCTGGCGGTG GCGGATCGGA TGTGTGATG ACTCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

TCTGGCGGTG GCGGATCGGA AATTGTGTTG ACGCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

26

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

TCTGGCGGTG GCGGATCGGA CATCGTGATG ACCCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

TCTGGCGGTG GCGGATCGGA AACGACACTC ACGCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

TCTGGCGGTG GCGGATCGGA AATTGTGCTG ACTCAGTCTC C

41

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid

27

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

TTCTCGTGCG GCCGCCTAAC GTTTGATTTC CACCTTGGTC CC

42

- (2) INFORMATION FOR SEQ ID NO: 60:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

TTCTCGTGCG GCCGCCTAAC GTTTGATCTC CAGCTTGGTC CC

42

- (2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

TTCTCGTGCG GCCGCCTAAC GTTTGATATC CACTTGGTC CC

42

- (2) INFORMATION FOR SEQ ID NO: 62:

28

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

TTCTCGTGCG GCCGCCTAAC GTTTGATCTC CACCTTGGTC CC

42

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

TTCTCGTGCG GCCGCCTAAC GTTTAATCTC CAGTCGTGTC CC

42

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 41 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

29

TCTGGCGGTG GCGGATCGCA GTCTGTGTTG ACGCAGCCGC C  
41

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

TCTGGCGGTG GCGGATCGCA GTCTGCCCTG ACTCAGCCTG C  
41

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthesis DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

TCTGGCGGTG GCGGATCGTC CTATGTGCTG ACTCAGCCAC C  
41

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthesis DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

TCTGGCGGTG GCGGATCGTC TTCTGAGCTG ACTCAGGACC C

41

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthesis DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

TCTGGCGGTG GCGGATCGCA CGTTATACTG ACTCAACCGC C

41

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthesis DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

TCTGGCGGTG GCGGATCGCA GGCTGTGCTC ACTCAGCCGT C

41

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthesis DNA"



31

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

TCTGGCGGTG GCGGATCGAA TTTTATGCTG ACTCAGCCCC A

41

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

TTCTCGTGCG GCCGCCTAAC CTAGGACGGT GACCTTGGTC CC

42

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

TTCTCGTGCG GCCGCCTAAC CTAGGACGGT CAGCTTGGTC CC

42

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

32

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

TTCTCGTGCG GCCGCCTAAC CTAAAACGGT GAGCTGGGTC CC

42

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

CGATCCGCCA CCGCCAGA

18

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "synthetic DNA"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

GTCTCCTCAG GTGGAGGC

18

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

33

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "synthetic DNA"

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

CGATCCGCCA CCGCCAGAGC CACCTCCGCC TGAACCGCCT CCACCTGAGG AGAC

54

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/GB 98/02630

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/10 C12N15/62 C07K14/72 C07K16/00  
C12Q1/68 G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 93 08278 A (AFFYMAX TECH NV) 29 April 1993 see page 9, line 32 - line 37 ---	1-23
A	US 5 597 693 A (EVANS RONALD M ET AL) 28 January 1997 see column 1, line 38 - line 49 see column 1, line 56 - line 61 ---	1-23
A	M. BROWN AND P.A. SHARP: "Human estrogen receptor forms multiple protein-DNA complexes" J. BIOL. CHEM., vol. 265, no. 19, 5 July 1990, pages 11238-11243, XP002092292 AM. SOC. BIOCHEM. MOL. BIOL., INC., BALTIMORE, US see the whole document --- -/--	1-23

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "3" document member of the same patent family

Date of the actual completion of the international search

4 February 1999

Date of mailing of the international search report

18/02/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx: 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

# INTERNATIONAL SEARCH REPORT

Ir. International Application No  
PC 98/02630

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 92 20791 A (CAMBRIDGE ANTIBODY TECH ;MEDICAL RES COUNCIL (GB)) 26 November 1992 cited in the application see the whole document ---	1-23
A	DE 196 46 372 C (EVOTEC BIOSYSTEMS GMBH) 19 June 1997 see the whole document ---	1-23
A	WO 92 01047 A (CAMBRIDGE ANTIBODY TECH ;MEDICAL RES COUNCIL (GB)) 23 January 1992 cited in the application see the whole document -----	1-23

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/GB 98/02630

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9308278	A	29-04-1993	US 5270170 A AU 3759693 A EP 0610448 A US 5498530 A US 5733731 A US 5338665 A	14-12-1993 21-05-1993 17-08-1994 12-03-1996 31-03-1998 16-08-1994
US 5597693	A	28-01-1997	AT 166360 T AU 655912 B AU 5342390 A CA 2047752 A DE 69032324 D DE 69032324 T EP 0463081 A JP 4505012 T WO 9011273 A	15-06-1998 19-01-1995 22-10-1990 18-09-1990 25-06-1998 08-10-1998 02-01-1992 03-09-1992 04-10-1990
WO 9220791	A	26-11-1992	AT 145237 T AU 664155 B AU 8221691 A CA 2086936 A DE 69123156 D DE 69123156 T DK 589877 T EP 0589877 A EP 0585287 A EP 0774511 A EP 0844306 A ES 2096655 T WO 9201047 A GR 3022126 T AU 665190 B AU 1693892 A CA 2109602 A JP 6508511 T AU 665025 B AU 2593392 A AU 665221 B AU 3089092 A AU 673515 B AU 3763893 A CA 2119930 A CA 2124460 A CA 2131151 A EP 0605522 A EP 0616640 A EP 0656941 A WO 9306213 A WO 9311236 A WO 9319172 A JP 6510671 T JP 7502167 T JP 7505055 T US 5565332 A US 5733743 A	15-11-1996 09-11-1995 04-02-1992 11-01-1992 19-12-1996 17-04-1997 07-04-1997 06-04-1994 09-03-1994 21-05-1997 27-05-1998 16-03-1997 23-01-1992 31-03-1997 21-12-1995 30-12-1992 26-11-1992 29-09-1994 14-12-1995 27-04-1993 21-12-1995 28-06-1993 14-11-1996 21-10-1993 01-04-1993 10-06-1993 30-09-1994 13-07-1994 28-09-1994 14-06-1995 01-04-1993 10-06-1993 30-09-1993 01-12-1994 09-03-1995 08-06-1995 15-10-1996 31-03-1998
DE 19646372	C	19-06-1997	NONE	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 98/02630

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9201047 A	23-01-1992	AT 145237 T	15-11-1996
		AU 664155 B	09-11-1995
		AU 8221691 A	04-02-1992
		CA 2086936 A	11-01-1992
		DE 69123156 D	19-12-1996
		DE 69123156 T	17-04-1997
		DK 589877 T	07-04-1997
		EP 0589877 A	06-04-1994
		EP 0585287 A	09-03-1994
		EP 0774511 A	21-05-1997
		EP 0844306 A	27-05-1998
		ES 2096655 T	16-03-1997
		WO 9220791 A	26-11-1992
		GR 3022126 T	31-03-1997
		AU 665190 B	21-12-1995
		AU 1693892 A	30-12-1992
		CA 2109602 A	26-11-1992
		JP 6508511 T	29-09-1994